

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:15:20 ; Search time 9505 Seconds
(without alignments)
10825.501 Million cell updates/sec

Title: us-09-105-117k-1

Perfect score: 2374

Sequence: 1 ccatttctgaaggtgttac.....ttcttccaaaggagtatct 2374

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hlg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_av.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_ats.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2374	100.0	2374	1	CGLYSEG	X96471 C-glutaminc
2	2374	100.0	2374	6	A93933	A93933 Sequence 2
3	2374	100.0	33150	1	AP005277	AP005277 Coryneb
4	2374	100.0	34980	6	AX127147	AX127147 Sequence
5	993	41.8	993	6	AX063767	AX063767 Sequence
6	993	41.8	993	6	AX244055	AX244055 Sequence
7	870	36.6	870	6	AX123540	AX123540 Sequence
8	870	36.6	870	6	BD165657	BD165657 Novel pol
9	822	34.6	822	6	AX063771	AX063771 Sequence
10	822	34.6	822	6	AX244059	AX244059 Sequence
11	818.6	34.5	308650	1	AP005218	AP005218 Coryneb
12	798.8	33.6	1771	1	AB083133	AB083133 Coryneb
13	724.4	30.5	349535	1	BX248357	BX248357 Coryneb
14	711	29.9	711	6	AX643028	AX643028 Sequence
15	708	29.8	708	6	AX123539	AX123539 Sequence
16	708	29.8	708	6	BD165656	BD165656 Novel pol
17	698.6	29.4	1568	6	E54483	E54483 Heat-resist
18	698.6	29.4	1568	6	AX643030	AX643030 Sequence
19	649	27.3	1083	6	AX813971	AX813971 Sequence
20	627	26.4	627	6	AX123538	AX123538 Sequence
21	627	26.4	627	6	BD165655	BD165655 Novel pol
22	485.2	20.4	993	6	AX067087	AX067087 Sequence
23	170	7.2	345783	1	AP003001	AP003001 Mesorhizo
24	164.4	6.9	8472	1	AE008076	AE008076 Agrobacte
25	164.4	6.9	13051	1	AE009111	AE009111 Agrobacte
26	159.4	6.7	302007	1	SC0939132	AL939132 Streptomy
27	140.6	5.9	303550	1	SC0939131	AL939131 Streptomy
28	138.6	5.7	306250	1	SME591786	AL591788 Sinorhizo
29	123.8	5.2	347356	1	BX640437	BX640437 Bordetell
30	123.8	5.2	348171	1	BX640412	U34849 Mycobacteri
31	122.8	5.2	14844	1	MBU34849	186263 Sequence 17
32	122.8	5.2	152339	6	I86263	BX248341 Mycobacte
33	122.2	5.1	306050	1	BX248341	AE007056 Mycobacte
34	121.2	5.1	348251	1	AE007056	AE007056 Mycobacte
35	121.2	5.1	348251	1	AE007056	AE007056 Mycobacte
36	121.2	5.1	348251	1	AE007056	AE007056 Mycobacte
37	120.8	5.1	298300	1	AP005025	AP005025 Streptomy
38	116.4	4.9	11120	1	RFA311775	AJ311775 Rhodococc
39	115.2	4.9	298300	1	AP005025	AP005025 Streptomy
40	113	4.8	303550	1	SC0939131	AL939131 Streptomy
41	112.6	4.7	10871	1	AE011965	AE011965 Xanthomon
42	112.4	4.7	10336	1	AE005530	AE005530 Escherich
43	112.2	4.7	12070	1	AB011413	AB011413 Streptomy
44	111.2	4.7	10278	1	AE015317	AE015317 Shigella
45	111.2	4.7	266658	1	AP002563	AP002563 Escherich

ALIGNMENTS

RESULT 1	CGLYSEG	2374 bp	DNA	linear	BCT 19-MAR-2001
LOCUS	C-glutaminc	lysE and lysG genes.			
DEFINITION	X96471				
ACCESSION	X96471.1	GI:1729753			
VERSION	lysE gene; lysG gene; Lysine export regulator protein; lysine				
KEYWORDS	exporter protein; lysine governor.				
SOURCE	Corynebacterium glutamicum				
ORGANISM	Corynebacterium glutamicum				
REFERENCE	Bacteria; Actinobacteria; Actinomycetales;				
AUTHORS	Corynebacterineae; Corynebacteriaceae; Corynebacterium.				
TITLE	Vrljic,M., Sahm,H. and Eggeling,L.				
	A new type of transporter with a new type of cellular function:				

JOURNAL L-lysine export from *Corynebacterium glutamicum*
 MEDLINE Mol. Microbiol. 22 (5), 815-826 (1996)
 PUBMED 97126810
 REFERENCE 8971704
 2 (bases 1 to 2374)
 AUTHORS Vrljic, M.M.
 TITLE Direct Submission
 JOURNAL Submitted (07-MAR-1996) M.M. Vrljic, Institut fuer Biotechnologie
 1, Forschungszentrum Juelich, Postfach 1913, D-52425 Juelich, FRG
 FEATURES Location/Qualifiers
 source
 1..2374
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /strain="R127"
 /db_xref="taxon:1718"
 complement (82..954)
 /gene="lysG"
 complement (82..954)
 /gene="lysG"
 /note="lysine governor"
 /codon_start=1
 /transl_table=11
 /product="lysine export regulator protein"
 /protein_id="CAA65323.1"
 /db_xref="GI:1729754"
 /db_xref="GOA:P94632"
 /db_xref="SWISS-PROT:P94632"
 /translation="MNPIQLTLLSISIDRGSFEGASIALSISPSAVSRVKALEHRVG
 RVLVSRTPAKATEAGEVIVQAKRMVLQAKTQALSGRLAEIPLATINADLSLSTW
 FPPVNEVASGGATITLRLDEAHTLSLRGQVGLGAVTREANFVAGCEVLELGTMR
 HLAITPSLRDYMVDGLDWAAMPVLRGPKVQLDRLDGRVDGPGVRRRSVLPVS
 AEGREGEATRLGLGWLIPETQAAPMLKAGEVILLDEIPDITPMYQWRLESRLARL
 TDVVDNAIEGLRP"
 1025..1726
 /gene="lysE"
 1025..1726
 /gene="lysE"
 /codon_start=1
 /transl_table=11
 /product="lysine exporter protein"
 /protein_id="CAA65324.2"
 /db_xref="GI:13397387"
 /db_xref="GOA:P94633"
 /db_xref="SWISS-PROT:P94633"
 /translation="MEIFITGLIGSILISIGPONVLVIKQIKRGLIAVLVCLLI
 SDVLEIFAGTLGVDLSNAEPIVLDIMRGGIAYLLFWAKAKDMINKVRFQIIE
 ETSPTVPDPTPLGSAAVATDTRNRVRVESVDKQKRVWKPMALVITVLTNENAYLDA
 FVFIGVGAQYGGDTGRNIPFAGAFPAASLIWFPLVGFAGAAALSRPLSSPKVWRINVVV
 AVMTALAIAKMLMG"
 complement (1723..2352)
 /gene="orf3"
 complement (1723..2352)
 /gene="orf3"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA65325.1"
 /db_xref="GI:1729756"
 /db_xref="SPRERE:R94634"
 /translation="NYALRDIVASGKALVYGISYSGPGLTABAEEFWAEGCPPLIHQ
 PYSIINRWPEPDGDDENLQSAANNGLGVIAFSPLAQSLTLDKLDLPGIPEGRASQ
 KSLSGMLNINNDVYKXINDIAQERGQSLAQWALVLRQEGYGDVITVTSALIGA
 SSVEQLDNLSDLSLNLEFSDAELEAIDEISHDAGINIWAKATDSKTREN"
 ORIGIN
 Query Match 100.0%; Score 2374; DB 1; Length 2374;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCATTTCGTGAAGGTGTTACTCTGCTGCTGCCCAATTCCTGCGGCGGAAGAGTGAATAAC 60
 Db 1 CCATTTCGTGAAGGTGTTACTCTGCTGCTGCCCAATTCCTGCGGCGGAAGAGTGAATAAC 60
 QY 61 CCTGAACCTTTTCAGAAAGTAACCTAAGGCGCGCAATCCCTCGATTGCTGCATCAACGCGC 120

Db 61 CCTGAACCTTTTCAGAAAGTAACCTAAGGCGCGCAATCCCTCGATTGCTGCATCAACGCGC 120
 QY 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTTCCAGGCGCCATCTGTGGCCAAATACATCGGTGT 180
 Db 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTTCCAGGCGCCATCTGTGGCCAAATACATCGGTGT 180
 QY 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTAGCATGGAGGAGCTTG 240
 Db 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTAGCATGGAGGAGCTTG 240
 QY 241 GGTTCGGGAAGAGTCCCAACCAAGGCGCTCGCGCAATTCGCTCACCAAAACCTTTCGCG 300
 Db 241 GGTTCGGGAAGAGTCCCAACCAAGGCGCTCGCGCAATTCGCTCACCAAAACCTTTCGCG 300
 QY 301 CGACGGGCAATGGATAGCGCTCGCGCCCAAGGACCATCGACGCGCGCTGCGAGGTC 360
 Db 301 CGACGGGCAATGGATAGCGCTCGCGCCCAAGGACCATCGACGCGCGCTGCGAGGTC 360
 QY 361 AGGCTCTGAGCACATCTTTGGGCGGAGGCTAAGCGGGCATCGCAGCCCAATCTAG 420
 Db 361 AGGCTCTGAGCACATCTTTGGGCGGAGGCTAAGCGGGCATCGCAGCCCAATCTAG 420
 QY 421 TTTTCCATCAACCATAGTAGCATCCCGCAATGAGGGGTTGCAATGSCCAAGTGGCGAT 480
 Db 421 TTTTCCATCAACCATAGTAGCATCCCGCAATGAGGGGTTGCAATGSCCAAGTGGCGAT 480
 QY 481 GGTTCGAAGTTCTACTTACTTCAATCCCGCACCGGATAGCTTCAAGGCTTACCGCTCC 540
 Db 481 GGTTCGAAGTTCTACTTACTTCAATCCCGCACCGGATAGCTTCAAGGCTTACCGCTCC 540
 QY 541 TAAACATCTCCAGCGCCGACGAGGATTAATGTGTGCGCTTCACTTCCAGGCGAGGCT 600
 Db 541 TAAACATCTCCAGCGCGACGAGGATTAATGTGTGCGCTTCACTTCCAGGCGAGGCT 600
 QY 601 GAGCGTTCCTCCACCCCAAGAGCTACTCTGTGTAACAGGGAGGAAACATGTGGATAG 660
 Db 601 GAGCGTTCCTCCACCCCAAGAGCTACTCTGTGTAACAGGGAGGAAACATGTGGATAG 660
 QY 661 CGAATCTCGGTTCATGCGGATGTTAAACGGGATTTCAAGGAGGCTCCAGATAGTTGCGC 720
 Db 661 CGAATCTCGGTTCATGCGGATGTTAAACGGGATTTCAAGGAGGCTCCAGATAGTTGCGC 720
 QY 721 TTTAGTTTCTGCTGACGACACCAATTTTCCGCGCTGCTTGCACAGGACTTCCACCGC 780
 Db 721 TTTAGTTTCTGCTGACGACACCAATTTTCCGCGCTGCTTGCACAGGACTTCCACCGC 780
 QY 781 TTGCGTTCCTTTGCGGTTGGGTCGATACCAACATCTCGACCCACGCTGATGCTCGAG 840
 Db 781 TTGCGTTCCTTTGCGGTTGGGTCGATACCAACATCTCGACCCACGCTGATGCTCGAG 840
 QY 841 AGCTTTAACGCGCTGACTCAACCGCGAGGGGAAATGAAAGGCTAAGGAGGCGGCTTC 900
 Db 841 AGCTTTAACGCGCTGACTCAACCGCGAGGGGAAATGAAAGGCTAAGGAGGCGGCTTC 900
 QY 901 GAGCTGCTTCTCATGATGATGAGGCAAGTCTCCAGTTCATGCGGTTCAATGAAGCT 960
 Db 901 GAGCTGCTTCTCATGATGATGAGGCAAGTCTCCAGTTCATGCGGTTCAATGAAGCT 960
 QY 961 ATATTAAACCATGTTAAGAACCAATCTTTAATTAAGTACTTCCATAGGTACAGATGGT 1020
 Db 961 ATATTAAACCATGTTAAGAACCAATCTTTAATTAAGTACTTCCATAGGTACAGATGGT 1020
 QY 1021 GATCATGGAATCTTCATTAACAGGTCTGCTTTTGGGGGCGAGTCTTTTACTGTCCATCGG 1080
 Db 1021 GATCATGGAATCTTCATTAACAGGTCTGCTTTTGGGGGCGAGTCTTTTACTGTCCATCGG 1080
 QY 1081 ACCCGAGAGTACTGGTGATTAACAGAGATTAAGCGGAGGACTCATTCCGGTCT 1140
 Db 1081 ACCCGAGAGTACTGGTGATTAACAGAGATTAAGCGGAGGACTCATTCCGGTCT 1140
 QY 1141 TCTCGTGTGTTAAATTTCTGAGCTGTTTTGTTCATCGCGGACCTTGGCGGTGATCT 1200

1141 TCTCGTGTGTTAAATTTCTGACGTCCTTTTGTCTCATGCCGCACTTGGGGGTGATCT 1200
 1201 TTTGTCAATGCGCGCGATCGTTCGATATATGCGTGGGGTGGCATCGCTTACCT 1260
 1201 TTTGTCAATGCGCGCGATCGTTCGATATATGCGTGGGGTGGCATCGCTTACCT 1260
 1261 GTTATGTTTGGCGTTCATGCGAGGAAAGACGCGATGACAAACAGGTGGAAGCGCAC 1320
 1261 GTTATGTTTGGCGTTCATGCGAGGAAAGACGCGATGACAAACAGGTGGAAGCGCAC 1320
 1321 GATCATTTGAAGAAACAGAACCAACCGTGGCGGATGACACGCTTTGGGGGTTCGGGCT 1380
 1321 GATCATTTGAAGAAACAGAACCAACCGTGGCGGATGACACGCTTTGGGGGTTCGGGCT 1380
 1381 GCGCATTCACACGCGCAACCGGGTGGGGTGGAGGTGAGCGTCATTAAGCAGCGGTTTG 1440
 1381 GCGCATTCACACGCGCAACCGGGTGGGGTGGAGGTGAGCGTCATTAAGCAGCGGTTTG 1440
 1441 GGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTGAAACCGCAATGCGTATTGGA 1500
 1441 GGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTGAAACCGCAATGCGTATTGGA 1500
 1501 CGCGTTGTGTTATGCGCGGCTCGCGCGCAATACGCGCACACCGCGGTGATTTT 1560
 1501 CGCGTTGTGTTATGCGCGGCTCGCGCGCAATACGCGCACACCGCGGTGATTTT 1560
 1561 CGCGCTGCGCGGTTGCGCGCAAGCTGATCTGCTTCCCGTGGTGGGTTTGGCGCAGC 1620
 1561 CGCGCTGCGCGGTTGCGCGCAAGCTGATCTGCTTCCCGTGGTGGGTTTGGCGCAGC 1620
 1621 AGCATTTGTCAGCGCGCTGTCAGCGCCCAAGGTGTGGCGTGGATCAACGTCGTGTCG 1680
 1621 AGCATTTGTCAGCGCGCTGTCAGCGCCCAAGGTGTGGCGTGGATCAACGTCGTGTCG 1680
 1681 AGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAGTTTTCGGGGTTT 1740
 1741 GGAATCGGTGGCTTCCGCCAATGTTGATCGCGGCTGTGGGAAATCTCATCGATCGC 1800
 1741 GGAATCGGTGGCTTCCGCCAATGTTGATCGCGGCTGTGGGAAATCTCATCGATCGC 1800
 1801 CTCCAACTCGCGCTCAGAAATCTCAAGTTGTGATGATCAAGGCTGTGTCAGCTG 1860
 1801 CTCCAACTCGCGCTCAGAAATCTCAAGTTGTGATGATCAAGGCTGTGTCAGCTG 1860
 1861 CTCACTGACGAGCAATCAATCAATGCTGCTGACGCTGATCGCGGCTGCTCTCTTG 1920
 1861 CTCACTGACGAGCAATCAATCAATGCTGCTGACGCTGATCGCGGCTGCTCTCTTG 1920
 1921 CTGCGCAGCAGCAGCAATCAATCAATGCTGCTGACGCTGATCGCGGCTGCTCTCTTG 1980
 1921 CTGCGCAGCAGCAGCAATCAATCAATGCTGCTGACGCTGATCGCGGCTGCTCTCTTG 1980
 1981 GTCAATGAGCTTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2040
 1981 GTCAATGAGCTTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2040
 2041 CTTCACCTGCGTGGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2100
 2041 CTTCACCTGCGTGGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2100
 2101 GCGCTGCGCAGTGGTGAAGAGCAATGACGCGCAAGCAGCAGCAGCAGCAGCAGCAG 2160
 2101 GCGCTGCGCAGTGGTGAAGAGCAATGACGCGCAAGCAGCAGCAGCAGCAGCAGCAG 2160
 2161 CAAGTTCTCAACGCTGATCGCGGCTTCTCCACCAAGCAGCAGCAGCAGCAGCAG 2220
 2161 CAAGTTCTCAACGCTGATCGCGGCTTCTCCACCAAGCAGCAGCAGCAGCAGCAG 2220
 2221 CTGATGATCAGAGCGGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2280
 2221 CTGATGATCAGAGCGGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2280

2281 CTTCTGAGCCGTAGGAGAAATACCCAGTAAGAGCTTTCCAGACGCAACATGTCAGC 2340
 2281 CTTCTGAGCCGTAGGAGAAATACCCAGTAAGAGCTTTCCAGACGCAACATGTCAGC 2340
 2341 CAATCGTACATGTTCTTCTCCAAAGGAGTATCT 2374
 2341 CAATCGTACATGTTCTTCTCCAAAGGAGTATCT 2374

RESULT 2
 A93933 2374 bp DNA linear PAT 22-JAN-2000
 LOCUS Sequence 2 from Patent WO9723597.
 DEFINITION A93933
 ACCESSION A93933
 VERSION A93933.1 GI:6742037
 KEYWORDS Corynebacterium glutamicum
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1 (bases 1 to 2374)
 AUTHORS Vrljic, M. and Eggeling, L.
 TITLE PROCESS FOR THE MICROBIAL PRODUCTION OF AMINO ACIDS BY BOOSTED
 ACTIVITY OF EXPORT CARRIERS
 JOURNAL Patent: WO 9723597-A 2 03-JUL-1997;
 KERNFORSCHUNGSANLAGE JUELICH (DE); VRLJIC MARINA (DE)

FEATURES
 source 1. .2374
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"

ORIGIN

Query Match 100.0%; Score 2374; DB 6; Length 2374;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATTTGCTGAGGTTTACTCTGCTGCGCCCAATCTCTGCGGGGAGAAAGTGAAGAAC 60
 Db 1 CCATTTGCTGAGGTTTACTCTGCTGCGCCCAATCTCTGCGGGGAGAAAGTGAAGAAC 60
 Qy 61 CCTGAACCTTTTCAAGAGTAACCTAAGCGCGCAATCCCTCGATTTGCTGCAATCAAGCGGC 120
 Db 61 CCTGAACCTTTTCAAGAGTAACCTAAGCGCGCAATCCCTCGATTTGCTGCAATCAAGCGGC 120
 Qy 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTCAGCGCGCATCTGTCCTGATTCATCATCGGTGT 180
 Db 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTCAGCGCGCATCTGTCCTGATTCATCATCGGTGT 180
 Qy 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTAGCATGGGAGAGCTTG 240
 Db 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTAGCATGGGAGAGCTTG 240
 Qy 241 GGTTCGCGGAGAGTGTCCCAACCAAGGCTCTGGGGAATTCCTCTACCAAAACCTTCGCG 300
 Db 241 GGTTCGCGGAGAGTGTCCCAACCAAGGCTCTGGGGAATTCCTCTACCAAAACCTTCGCG 300
 Qy 301 CGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Db 301 CGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Qy 361 AGGTCCTTGAAGCAATCTTTGGGACCGAAGGCTGAAGCGGATTAAGCGGATTAAGCGGAT 420
 Db 361 AGGTCCTTGAAGCAATCTTTGGGACCGAAGGCTGAAGCGGATTAAGCGGATTAAGCGGAT 420
 Qy 421 TTTTCCATCAACCATGATGAGGATCCCGCAATGAGGGGTTCGATGGCCAGTGGCGCAT 480
 Db 421 TTTTCCATCAACCATGATGAGGATCCCGCAATGAGGGGTTCGATGGCCAGTGGCGCAT 480
 Qy 481 GGTTCGAAGTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 540
 Db 481 GGTTCGAAGTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 540

QY	541	TAATAACATCTCCAGCGCCGAGAGGATAATGTGTGCGCTTCATCTTCCAGCGCAGCGT	600
Db	541	TAATAACATCTCCAGCGCCGAGAGGATAATGTGTGCGCTTCATCTTCCAGCGCAGCGT	600
QY	601	GAGCGTTGCTCCACCCCAAGAGCTACCTCGTTGAACACGCGGAGAAACCATGTGGATAG	660
Db	601	GAGCGTTGCTCCACCCCAAGAGCTACCTCGTTGAACACGCGGAGAAACCATGTGGATAG	660
QY	661	CGAATCTGCGCTGATGCGATGTTAAAGGATTTTCAAGCGGCTCCAGATAGTTGCGC	720
Db	661	CGAATCTGCGCTGATGCGATGTTAAAGGATTTTCAAGCGGCTCCAGATAGTTGCGC	720
QY	721	TTTAGTTCTGCTTCAGCAACACCATTTTCGCGCTGCTTGGCAGAGACTTTACCGCG	780
Db	721	TTTAGTTCTGCTTCAGCAACACCATTTTCGCGCTGCTTGGCAGAGACTTTACCGCG	780
QY	781	TTTCGGTTGCTTTGGCGGCTTGGGTGCGGATACCAACACTCGACCCACGTTGATGTCGAG	840
Db	781	TTTCGGTTGCTTTGGCGGCTTGGGTGCGGATACCAACACTCGACCCACGTTGATGTCGAG	840
QY	841	AGCTTTAAGCGCTGACTCACCGCGAGGGGAAATGGAAGGGCTAAGGAGGCGCTTC	900
Db	841	AGCTTTAAGCGCTGACTCACCGCGAGGGGAAATGGAAGGGCTAAGGAGGCGCTTC	900
QY	901	GAGCGCTCTCATCAATGATGAGAGCAAGTGTCCAGTTGATGGGGTTCATGAGCT	960
Db	901	GAGCGCTCTCATCAATGATGAGAGCAAGTGTCCAGTTGATGGGGTTCATGAGCT	960
QY	961	ATATTAACCACTGTTAAGAACCAATCATTTTACTTAAGTACTTCCATAGTCCAGTGT	1020
Db	961	ATATTAACCACTGTTAAGAACCAATCATTTTACTTAAGTACTTCCATAGTCCAGTGT	1020
QY	1021	GATCATGGAATCTTATTAACAGTCTGCTTTTGGGGGAGTCTTTTACTGTCCTCAGG	1080
Db	1021	GATCATGGAATCTTATTAACAGTCTGCTTTTGGGGGAGTCTTTTACTGTCCTCAGG	1080
QY	1081	ACCGCAGAACTGATGTTGATTAACAGGAAATTAAGCGCGGAGAACTCATTTGCGGTTCT	1140
Db	1081	ACCGCAGAACTGATGTTGATTAACAGGAAATTAAGCGCGGAGAACTCATTTGCGGTTCT	1140
QY	1141	TTCTGTTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCGGACCTTTGGGTTGATCT	1200
Db	1141	TTCTGTTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCGGACCTTTGGGTTGATCT	1200
QY	1201	TTTGTCCAAATGCGCGCGCTGCTGCTCGATATTATGCTGCTGGGTGCTGCTTACCT	1260
Db	1201	TTTGTCCAAATGCGCGCGCTGCTGCTCGATATTATGCTGCTGGGTGCTGCTTACCT	1260
QY	1261	GTTATGTTTGGCGTCATGGCAGAAAGACCGCCATGACAAAACAGGTGGAAGCGCCACA	1320
Db	1261	GTTATGTTTGGCGTCATGGCAGAAAGACCGCCATGACAAAACAGGTGGAAGCGCCACA	1320
QY	1321	GATCATGGAAGAACAGAACCAACCGTGGCGGATGACACGCTTTGGCGGTTTCGGCGGT	1380
Db	1321	GATCATGGAAGAACAGAACCAACCGTGGCGGATGACACGCTTTGGCGGTTTCGGCGGT	1380
QY	1381	GCGCATGACACCGCGACCGGTTGCGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG	1440
Db	1381	GCGCATGACACCGCGACCGGTTGCGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG	1440
QY	1441	GCTAAGCCCATGTTGATGCAATCGTGTGCTGACTGTTGAAACCCGATCGTATTGGA	1500
Db	1441	GCTAAGCCCATGTTGATGCAATCGTGTGCTGACTGTTGAAACCCGATCGTATTGGA	1500
QY	1501	CGCGTTGTTGTTATCGCGCGCTCGCGCGCAATACGCGACACCGAGCGGTGATTTT	1560
Db	1501	CGCGTTGTTGTTATCGCGCGCTCGCGCGCAATACGCGACACCGAGCGGTGATTTT	1560
QY	1561	CGCGCTGCGCGCTTCGCGCAAGCCTGATCTGTTCCGCTGGTGGGTTTCGCGCGAGC	1620
Db	1561	CGCGCTGCGCGCTTCGCGCAAGCCTGATCTGTTCCGCTGGTGGGTTTCGCGCGAGC	1620
QY	1621	ASCATTTGTCAGCGCCGCTGCTCCAGCCCAAGGTGTGCGCTGATCAACACGCTGCTGCGC	1680
Db	1621	ASCATTTGTCAGCGCCGCTGCTCCAGCCCAAGGTGTGCGCTGATCAACACGCTGCTGCGC	1680
QY	1681	AGTTGTGATGACCGAATGCGCATCAAACTGATGTATGAGTTAGTTTTCGCGGGTTT	1740
Db	1681	AGTTGTGATGACCGAATGCGCATCAAACTGATGTATGAGTTAGTTTTCGCGGGTTT	1740
QY	1741	GGAATCGTGGCTTTCGCGCAAAATGTTGATGCGCGCTGCGGAAATCTCATCGATCGC	1800
Db	1741	GGAATCGTGGCTTTCGCGCAAAATGTTGATGCGCGCTGCGGAAATCTCATCGATCGC	1800
QY	1801	CTCCACTGCGCTGTCAGAAATCTCCAAAGTTGTTGATGCAATCAAGGCTGTTGTCAGCTG	1860
Db	1801	CTCCACTGCGCTGTCAGAAATCTCCAAAGTTGTTGATGCAATCAAGGCTGTTGTCAGCTG	1860
QY	1861	CTCACTGACGAGACCAATCAATGACATGCTGCGCAAGTGAATCTGCGCGCTTCTGCGCGAT	1920
Db	1861	CTCACTGACGAGACCAATCAATGACATGCTGCGCAAGTGAATCTGCGCGCTTCTGCGCGAT	1920
QY	1921	CTCGCGAGACCCATGCAAGCGCATCTGCGCAAGTGAATCTGCGCGCTTCTGCGCGAT	1980
Db	1921	CTCGCGAGACCCATGCAAGCGCATCTGCGCAAGTGAATCTGCGCGCTTCTGCGCGAT	1980
QY	1981	GTCAATGAGCTTGGCGGACCATATCAATATTGTTCACTGTTCAACATGCCCTCAGACAGGA	2040
Db	1981	GTCAATGAGCTTGGCGGACCATATCAATATTGTTCACTGTTCAACATGCCCTCAGACAGGA	2040
QY	2041	CTTACCTGCTTGGCGGGAACCTCTGGAATTCATTCAGATATTTTGGCGAGCTGCTGCAA	2100
Db	2041	CTTACCTGCTTGGCGGGAACCTCTGGAATTCATTCAGATATTTTGGCGAGCTGCTGCAA	2100
QY	2101	GCCTTCGCGCAAGTGTGAGAAAGCAATGACGCAAGACCAATTTGTCGAGCTGACTGCAA	2160
Db	2101	GCCTTCGCGCAAGTGTGAGAAAGCAATGACGCAAGACCAATTTGTCGAGCTGACTGCAA	2160
QY	2161	CAAGTTCTCAGCGCTACCGCGCTTCTCCACCAAGATTAATGATGATGATGCTTGG	2220
Db	2161	CAAGTTCTCAGCGCTACCGCGCTTCTCCACCAAGATTAATGATGATGATGCTTGG	2220
QY	2221	CTGATGAATGAGAGCGGCGCTTCTCCGCAATGAACTCAGCGCGCTCGCTGAG	2280
Db	2221	CTGATGAATGAGAGCGGCGCTTCTCCGCAATGAACTCAGCGCGCTCGCTGAG	2280
QY	2281	CTCTGACCGTAGAGAAATACCCAGTAAGAGCTTTCCAGAGCAACAATGTCAAG	2340
Db	2281	CTCTGACCGTAGAGAAATACCCAGTAAGAGCTTTCCAGAGCAACAATGTCAAG	2340
QY	2341	CAATGCTACATGTTTCTTCCAAAGGATATCT	2374
Db	2341	CAATGCTACATGTTTCTTCCAAAGGATATCT	2374
RESULT 3			
AP005277/c			
LOCUS			
DEFINITION			
AP005277 333150 bp DNA linear BCT 08-AUG-2002			
Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section			
4/10			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Corynebacterium glutamicum ATCC 13032			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
Corynebacteriaceae; Corynebacterium.			
1			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.			

Thu Mar 18 12:30:58 2004

Ird., Tokyo Research Laboratories; 3-6-6, Asahi-nachi, Machida,
Tokyo 194-8533, Japan (E-mail: snakagawa@xanagen.com,
Tel: 81-44-829-3031, Fax: 81-44-813-1651)
This sequence is conducted by collaboration of Kyowa Hakko Kogyo
Co. Ltd. and Kitaaso University.

FEATURES

Location/Qualifiers
1..333150
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:196627"
/note="ATCC 13032"
complement(63..1292)
/gene="Cg10944"
complement(63..1292)
/gene="Cg10944"
/codon_start=1
/transl_table=11
/product="Hypothetical membrane protein"
/protein_id="BAB98337.1"
/db_xref="GI:21323711"
/translation="MTNPTTEARNRLIWLQNGIDQIVAAKTVLPMLQAQAPG
FILALLPIREAGSMPLQQAITGWLITQTSKVMVIGSQFVSALGIGVLAFLRG
WALGITVLIALSLPRSCMSIAKSVQKGLVTRATVIGVWGLVAGL
ATAFLGSHSPTRVLAANVTASSPMFASIVARIEYAKPATPKNAPSANPWRRCI
AALKDKAFRRFVLVRSMLVTALSTAFILVLAESGNSIDSLGFLIASGLASWVG
RISGWSHSSKNWAGALFGSIVLILVLSAPAPAIQNTLVPLSPFLILTAHTA
IRVARTYVMQAGDQRTYRVADANTLMVLLVLAUSGFIALFGNEALLFLAAI
GLIGTISARGLEKVSAG"
1382..2917
/gene="Cg10945"
1382..2917
/gene="Cg10945"
/codon_start=1
/transl_table=11
/product="Putative multicopper oxidases"
/protein_id="BAB98338.1"
/db_xref="GI:21323712"
/translation="MPELNRRTFFKGAVLAATVGAQVILVACSSDDVGVGGEPRTL
PIPPADLTREGSVHFALAEQSGQILPDVTKTWGFNGHLGFTLVVKGDDVHV
DVNLNDEMTVHMGKLPALADGPHSPGQGTWSTVFWANDAAATLHYPHTHG
LTGLHVRGLAGMIIVDEADTKDLDPREYGVDDIPLVMDHRLFEDGSLDEBDPLD
GLIGDPTANGITNAHPDATTRVRFRVLGNSNMRFNLAFTDRTFQVIAQSGLLD
EPQRTTLAIGCRKEWIVELPEGSDVTLSSGVFDYGVDPDEFVDFQMSQSPQL
LTITGPSDAAPALPGVLVKSTPDPDIDATERFTIMNTFESINDLQMDQORVDVWD
HDQPEVIVTNDNSDPNPFHDARFKVLKPEGDVLEFNGDKDTVGLPPGATIL
AVEFGHPDPQWPMYMYHCHMLYHEDQGMNGQFVIVPEGDEPAVVGSGTSSIDSAGG
HAH"
3063..3983
/gene="Cg10946"
3063..3983
/gene="Cg10946"
/note="PF00005:ABC transporter
TIGR00960:3a0501s02: Type II (General) Secretory Pathway
(IISP) Family protein
TIGR00968:3a0106s01: sulfate transport system permease
protein
TIGR01166:cbiO: cobalt transport protein ATP-binding
subunit
TIGR01184:ntrCD: nitrate transport ATP-binding subunits C
and D
TIGR01186:proV: glycine betaine/L-proline transport ATP
binding subunit
TIGR01187:potA: spermidine/putrescine ABC transporter
ATP-binding subunit
TIGR01188:draA: daunorubicin resistance ABC transporter
ATP-binding subunit
TIGR01189:ccmA: heme exporter protein CcmA"
/codon_start=1
/transl_table=11
/product="ABC-type transporter, ATPase component"
/protein_id="BAB98339.1"
/db_xref="GI:21323713"

/translation="MTTTPAIDVTIDYRTGYTAVKGLNEHVORGEVGLLGTNGAG
KSTLVEIGLSVSGTVRISGLDVPADRAILRIBELGLMLOSGLPSQLTVAETMDM
WHTCTVPRAIKDVLADVDLHRENVGALSGORLIDACALIGDPSILFLDEPT
TGLDESRRTWQLLDLKQGVMTLTHLEAEAFCLCDRIAIMNAGEIAGEVILDE
LVAREKSTISFVLGGQVELPVLGABIIIRDNNHVRITATTTLQOHTLEILTWAAETGI
ALEGFAKATLSEVFMDSIENTSLQTA"
4003..4836
/gene="Cg10947"
4003..4836
/gene="Cg10947"
/codon_start=1
/transl_table=11
/product="ABC-type transporter, permease components"
/protein_id="BAB98340.1"
/db_xref="GI:21323714"
/translation="MTTSHARGLOHATPERRKTSFFKTSFKAMLOFRNKTLIFM
ATVFGVPIPLLLFLINGGAASANSFDYFVMTLLFVQFVYSMAITRRDRSVLKR
LRTGARIDIIGAIKCPGALLTLFTVVIPIPLMVLGAPINLPIVFAVLIGLLL
CSALAMSGTFRNAEAQMTSPVLMAGGLGSRFVGDSIVADILATPTFAAIS
DLVQIGWAGTADSVGVGEARAFAGIQDMLIPGLILAATAAANAANRYMRWDSY
R"
4879..6138
/gene="Cg10948"
4879..6138
/gene="Cg10948"
/EC_number="2.7.3.-"
/codon_start=1
/transl_table=11
/product="Two-component system, sensory transduction
histidine kinases"
/protein_id="BAB98341.1"
/db_xref="GI:21323715"
/translation="MKNDFWAGTARWFSRGVSLASPVTAPLNSRRRLNLAKYTL
YTRVSLQAIPIVVLISLAFGLGVANAGTLNPFVWLLGFSVILLIVTVLVSEVQSLNS
HPRSVQPPFTGLVNLGVVSVVQLPGLNMSDNTRATLFTLCVFLLSIAYI
PMNRYWMLIAMSALVMTSTTDTLSALWVPIPLMAGTVLSVTVTVVWKEVERK
RELEASLVTERLFAQLHDLTGLQHLAAMSVKSELAKRGGDRLENELELOK
LRTSMEGRDVSVGYRTVNLATEISGAKSLIADATHLSVIGTTSQVSPAHRELCAW
LVREATTNLRISDAIDATLTLSTSVRMNNGVNGKIGRLSGLSALRSASAGMTL
IVSRDDQFSVRMLINAPANTPAEKA"
6139..6750
/gene="Cg10949"
6139..6750
/gene="Cg10949"
/note="PF00072:Response regulator receiver domain
PF00196:Bacterial regulatory proteins, luxR family"
/codon_start=1
/transl_table=11
/product="Two-component system, response regulators
consisting of a CheY-like receiver domain and a HTH
DNA-binding domain"
/protein_id="BAB98342.1"
/db_xref="GI:21323716"
/translation="MISIAADDEALIASLATLISLEPDLDPVPTAGSGELIETWA
DPSNTDVCVLIDQLGGIDGTATRLMETTDLAVLIVTSHARPOLKRALAAGVIG
FLPKTSADFEATARTHAGRYIDPELAAMTISAGESPLTNRREEVLELAGQGLSA
EELIAVHAHAFGTRNYLSQAKTVKQAFTRARELQML"
complement(6781..6987)
/gene="Cg10950"
complement(6781..6987)
/gene="Cg10950"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB98343.1"
/db_xref="GI:21323717"
/translation="MKHYMSYFGAGPLNTPTHTKEYVFMFDIVSVNHFWEKSTNNWVN
WGLPRETFVPAFVGSTVLAVLSF"
7048..8205
/gene="Cg10951"
7048..8205
/gene="Cg10951"
/codon_start=1

FEATURES		Location/Qualifiers	
source		1..349980	
		/organism="Corynebacterium glutamicum"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:1718"	
		/note="Seq 1 to long (3.309.400) split in 11, seq 7063	
		1.200.001 1.549.980"	
ORIGIN		Query Match	
		100.0%; Score 2374; DB 6; Length 349980;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 2374; Conservative 0; Mismatches 0; Gaps 0;	
QY	1	CGATTTCCTGAGAGTGTACTCTGCTGCGCCCAATTCCTGCGGCGCAAGAGTGAAGAAC 60	
Db	129968	CCATTTCCTGAGAGTGTACTCTGCTGCGCCCAATTCCTGCGGCGCAAGAGTGAAGAAC 129909	
QY	61	CCTGAACCTTTTCAGAAAGTAACTTAAGCGCGCAATCCCTCGATGTGTCATCAAGCAGCGC 120	
Db	129908	CCTGAACCTTTTCAGAAAGTAACTTAAGCGCGCAATCCCTCGATGTGTCATCAAGCAGCGC 129849	
QY	121	GTCTGTGAGTGTAGCTAGAGATCTAGATTCAGAGCGCATCGTTGCCAATACATCGGTGT 180	
Db	129848	GTCTGTGAGTGTAGCTAGAGATCTAGATTCAGAGCGCATCGTTGCCAATACATCGGTGT 129789	
QY	181	GTCAATGGGTATCTCATCGAGAGATCACTTCTCTGCTTTTAGCATGGGAGCAGTTG 240	
Db	129788	GTCAATGGGTATCTCATCGAGAGATCACTTCTCTGCTTTTAGCATGGGAGCAGTTG 129729	
QY	241	GGTTTCGGGGAAGAAGTCCCAACCAAGGGCTCGGGAAATTCCTCACCACCAACCTTCGCG 300	
Db	129728	GGTTTCGGGGAAGAAGTCCCAACCAAGGGCTCGGGAAATTCCTCACCACCAACCTTCGCG 129669	
QY	301	CGAGCGGCAATGATGATGATCGCGCTCGCCCAAGAGCATTCAGACGCGCGCTCCAGGTC 360	
Db	129668	CGAGCGGCAATGATGATGATCGCGCTCGCCCAAGAGCATTCAGACGCGCGCTCCAGGTC 129609	
QY	361	ACGGTCTTGAAGCACATCTTTGGGACCGAAGCGTGAAGACGCGCATCGAGCGCATCTAG 420	
Db	129608	ACGGTCTTGAAGCACATCTTTGGGACCGAAGCGTGAAGACGCGCATCGAGCGCATCTAG 129549	
QY	421	TTTCCCATCAACCATGTAGGCATCCCGCATGAGGGGTTCGATGCGCAAGTGGCGCAT 480	
Db	129548	TTTCCCATCAACCATGTAGGCATCCCGCATGAGGGGTTCGATGCGCAAGTGGCGCAT 129489	
QY	481	GGTTCCAAAGTTCTACTTCTACTTCAATCCCGCATGAGGGGTTCGATGCGCAAGTGGCGCAT 540	
Db	129488	GGTTCCAAAGTTCTACTTCTACTTCAATCCCGCATGAGGGGTTCGATGCGCAAGTGGCGCAT 129429	
QY	541	TAAGACATCTCCACGCGCAGCAAGGATTAATGTGTGCGCTTCACTTCCAGCGCAGCGT 600	
Db	129428	TAAGACATCTCCACGCGCAGCAAGGATTAATGTGTGCGCTTCACTTCCAGCGCAGCGT 129369	
QY	601	GAGCGTGTCTCCACCAAGAGTACTCTGTTGACACGCGGAGGAAACCATGTGATAG 660	
Db	129368	GAGCGTGTCTCCACCAAGAGTACTCTGTTGACACGCGGAGGAAACCATGTGATAG 129309	
QY	661	CGAATCTGCGTTGATGCGGATGTTAAAGCGATTTTCAAGAGCGCTCCAGATAGTTGGCG 720	
Db	129308	CGAATCTGCGTTGATGCGGATGTTAAAGCGATTTTCAAGAGCGCTCCAGATAGTTGGCG 129249	
QY	721	TTTAGTCTTCTGCTGACAGCAACACATTTTCCGCGTGTCTGCAAGAGCTTCACCCCGC 780	
Db	129248	TTTAGTCTTCTGCTGACAGCAACACATTTTCCGCGTGTCTGCAAGAGCTTCACCCCGC 129189	
QY	781	TTTCGGTGTCTGCGGTTGGGTGGGATACCAACATCTGACACCGCATGTGATCTCGAG 840	
Db	129188	TTTCGGTGTCTGCGGTTGGGTGGGATACCAACATCTGACACCGCATGTGATCTCGAG 129129	
QY	841	AGCTTTAAGCGCTGACTCACCAGCGGCGGGAATGGAAGGGCTTAAGAGGGCGCTTC 900	
Db	129128	AGCTTTAAGCGCTGACTCACCAGCGGCGGGAATGGAAGGGCTTAAGAGGGCGCTTC 129069	

QY	901	GAAGCTGCTTCATCAATGATTTGAGACAAAGTTCAGTTGAATGGGGTTCATGAGCT	960
Db	129068	GAAGCTGCTTCATCAATGATTTGAGACAAAGTTCAGTTGAATGGGGTTCATGAGCT	129009
QY	961	ATATTAAACCATGTTAGAACCAATCAATTTTACTTAAGTACTTCCATAGGTCAAGATGGT	1020
Db	129008	ATATTAAACCATGTTAGAACCAATCAATTTTACTTAAGTACTTCCATAGGTCAAGATGGT	128949
QY	1021	GATCATGGAATCTTCAATTACAGTCTCTCTTTTGGGGCCAGTCTTTTACTGTCCATCGG	1080
Db	128948	GATCATGGAATCTTCAATTACAGTCTCTCTTTTGGGGCCAGTCTTTTACTGTCCATCGG	128889
QY	1081	ACCGCAGAATGACTGTGTGATTAACACAGGAATTTAGCGGGAAGGACTCAATCCGGTCTCT	1140
Db	128888	ACCGCAGAATGACTGTGTGATTAACACAGGAATTTAGCGGGAAGGACTCAATCCGGTCTCT	128829
QY	1141	TCTCGTGTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCCGGCACCTTTGGCGGTTGATCT	1200
Db	128828	TCTCGTGTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCCGGCACCTTTGGCGGTTGATCT	128769
QY	1201	TTTGTCCAATGCGCGCCGATGCTGTGCTCGATATATATGCTGTGGGTGGCATCGCTTACCT	1260
Db	128768	TTTGTCCAATGCGCGCCGATGCTGTGCTCGATATATATGCTGTGGGTGGCATCGCTTACCT	128709
QY	1261	GTTATGCTTTGCGCGTATGAGGAAAGACGCCATGACAAACAGAGTGGAAAGCGCAC	1320
Db	128708	GTTATGCTTTGCGCGTATGAGGAAAGACGCCATGACAAACAGAGTGGAAAGCGCAC	128649
QY	1321	GATCATTGAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGCGGTTGCGGGT	1380
Db	128648	GATCATTGAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGCGGTTGCGGGT	128589
QY	1381	GGCCACTGACACGCGCAACCGGGTTCGGGTGAGGTGAGCGTGGATAGACGACGGGTTG	1440
Db	128588	GGCCACTGACACGCGCAACCGGGTTCGGGTGAGGTGAGCGTGGATAGACGACGGGTTG	128529
QY	1441	GGTAAAGCCCATGTTGATGGCAATCGTCTGACTGTTGAACCCGAAATCGTATTTGGA	1500
Db	128528	GGTAAAGCCCATGTTGATGGCAATCGTCTGACTGTTGAACCCGAAATCGTATTTGGA	128469
QY	1501	CGCGTTGTGTTTATCGGGGGTTCGGCGGCAATACGGCGACACCGGACGGTGAATTT	1560
Db	128468	CGCGTTGTGTTTATCGGGGGTTCGGCGGCAATACGGCGACACCGGACGGTGAATTT	128409
QY	1561	CGCGCTGCGCGGTTTCGGCGCAAGCTGATCGTGTGCGGTTCGGCGGACG	1620
Db	128408	CGCGCTGCGCGGTTTCGGCGCAAGCTGATCGTGTGCGGTTCGGCGGACG	128349
QY	1621	AGCATTTGACCGCGCTGTCAGCCCAAGTGTGCGCTGATCAACGCTCGTGTGCG	1680
Db	128348	AGCATTTGACCGCGCTGTCAGCCCAAGTGTGCGCTGATCAACGCTCGTGTGCG	128289
QY	1681	AGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAGTTTTCGGGGTTT	1740
Db	128288	AGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAGTTTTCGGGGTTT	128229
QY	1741	GGAATTCGGTGGCTTCGCCCAATGTTGATGCGCGCTGTGGGAATCTCATCGATCG	1800
Db	128228	GGAATTCGGTGGCTTCGCCCAATGTTGATGCGCGCTGTGGGAATCTCATCGATCG	128169
QY	1801	CTTCAATCTGGCGTCAAGAACTCCAAGTTGTTGAGTGAATCAAGCTGTTGTTCCAGCTG	1860
Db	128168	CTTCAATCTGGCGTCAAGAACTCCAAGTTGTTGAGTGAATCAAGCTGTTGTTCCAGCTG	128109
QY	1861	CTCACTGACGAGACCAATCAATGACCTGTTACGGTATCCGCGCGTACTCTCTTTG	1920
Db	128108	CTCACTGACGAGACCAATCAATGACCTGTTACGGTATCCGCGCGTACTCTCTTTG	128049
QY	1921	CTCGCGACGACCAATGACCTGCGCAAGTGAATGCGCGGTTCTCTGGCGAT	1980
Db	128048	CTCGCGACGACCAATGACCTGCGCAAGTGAATGCGCGGTTCTCTGGCGAT	127989
QY	1981	GTCAATGAGCTTGGGACCAATCAATATTGTTTCACTGTTCAACATGCGCTCAGACAGGGA	2040

QY 662 GAATCTCGTTGATGCGATGCTTAACGGGATTTTACGAGCGTCCAGATAGTTGCGCT 721
 Db 393 GAACTCGCTTGATGCGATGCTTAACGGGATTTTACGAGCGTCCAGATAGTTGCGCT 334
 QY 722 TTAGTTTCTGCTTGAGCAGACACATTTTCCGCGCTGCTTGACAGGACTTCAACCGCT 781
 Db 333 TTAGTTTCTGCTTGAGCAGACACATTTTCCGCGCTGCTTGACAGGACTTCAACCGCT 274
 QY 782 TCGGTTCTGCTTGCGCGGTTGGTGGCGGATACCAACATCGACCGCATGATCTCGAGA 841
 Db 273 TCGGTTCTGCTTGCGCGGTTGGTGGCGGATACCAACATCGACCGCATGATCTCGAGA 214
 QY 842 GCTTTACGCGCTGACTCAGCGCGAGGGGGAATGGAAGGGCTTAGAGGGCGCTTCG 901
 Db 213 GCTTTACGCGCTGACTCAGCGCGAGGGGGAATGGAAGGGCTTAGAGGGCGCTTCG 154
 QY 902 AAGCTGCTTTCATCAATGATGAGAGCAAAAGTGTCCAGTTGAATGGGTTTCATGAAGCTA 961
 Db 153 AAGCTGCTTTCATCAATGATGAGAGCAAAAGTGTCCAGTTGAATGGGTTTCATGAAGCTA 94
 QY 962 TATTAAACCATGTTAAGAACCAATCTTTTACTTAACTACTTCCATAGTCCAGATGCGTG 1021
 Db 93 TATTAAACCATGTTAAGAACCAATCTTTTACTTAACTACTTCCATAGTCCAGATGCGTG 34
 QY 1022 ATCATGGAATCTTCATTACAGGTCGTCTTTTG 1054
 Db 33 ATCATGGAATCTTCATTACAGGTCGTCTTTTG 1
 RESULT 7
 AX123540/c
 LOCUS AX123540 870 bp DNA linear PAT 11-MAY-2001
 DEFINITION Sequence 3456 from Patent EPI108790.
 ACCESSION AX123540
 VERSION AX123540.1 GI:14041028
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotides
 JOURNAL Patent: EP 1108790-A 3456 20-JUN-2001;
 KYOWA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source Location/Qualifiers
 1..870
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 ORIGIN
 Query Match 36.6%; Score 870; DB 6; Length 870;
 Best Local Similarity 100.0%; Pred. No. 9.5e-203;
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 85 AGCGCGCAATCCCTGATGCTGCTCAACAGCGGCTCTGTGAGTCTAGCTAGATCT 144
 Db 870 AGCGCGCAATCCCTGATGCTGCTCAACAGCGGCTCTGTGAGTCTAGCTAGATCT 811
 QY 145 AGATTCAGGCGCATGCTGCTCCCAATACATCGGTGTGTCATAGGATCTCTCATCGAGGAG 204
 Db 810 AGATTCAGGCGCATGCTGCTCCCAATACATCGGTGTGTCATAGGATCTCTCATCGAGGAG 751
 QY 205 GATCACTTCTCTCTTTTACGATGCGGAGCAGTGGTGTGGGAAGAGTCCCAAC 264
 Db 750 GATCACTTCTCTCTTTTACGATGCGGAGCAGTGGTGTGGGAAGAGTCCCAAC 691
 QY 265 AAGGCTCGGCAATGCTTCCAAACCTTCCGCCGAGCGGACAATGAGATCGGCT 324
 Db 690 AAGGCTCGGCAATGCTTCCAAACCTTCCGCCGAGCGGACAATGAGATCGGCT 631

QY 325 GCGCCCCACAGGACCATCGAGCGGCCGCTCCAGGTACGCTCTTGAGACATCTTTGGG 384
 Db 630 GCGCCCCACAGGACCATCGAGCGGCCGCTCCAGGTACGCTCTTGAGACATCTTTGGG 571
 QY 385 ACCGAGCGTGAAGAGCGGCAATCGAGCGCAATCTAGTTTCCATCAACCATGAGGCAATC 444
 Db 570 ACCGAGCGTGAAGAGCGGCAATCGAGCGCAATCTAGTTTCCATCAACCATGAGGCAATC 511
 QY 445 CCGCAATAGAGGGGTTGCAATGGCCAAAGTGGCGCATGGTTCCAAAGTTCTACTACTTACA 504
 Db 510 CCGCAATAGAGGGGTTGCAATGGCCAAAGTGGCGCATGGTTCCAAAGTTCTACTACTTACA 451
 QY 505 TCCGCGCCACGGGATTTAGCTTACGGGTTACCGCTTCTAAACATCTCCAGCGCGCAGCAA 564
 Db 450 TCCGCGCCACGGGATTTAGCTTACGGGTTACCGCTTCTAAACATCTCCAGCGCGCAGCAA 391
 QY 565 GGATAATGTTGCGCTTTCATCTTCCAAAGCGAGCGGTGAGCGTTGCTCCACCCCAAGAAGC 624
 Db 390 GGATAATGTTGCGCTTTCATCTTCCAAAGCGAGCGGTGAGCGTTGCTCCACCCCAAGAAGC 331
 QY 625 TACCTCGCTGAACAGCGGAGGACCATGTGATAGCGATCTCGGTGATGCGGATGCT 684
 Db 330 TACCTCGCTGAACAGCGGAGGACCATGTGATAGCGATCTCGGTGATGCGGATGCT 271
 QY 685 TAAAGCGGATTCAGCAAGCGGCTCCAGATAGTTGCGCTTTAGTTTCTGCTTGCAGCAAC 744
 Db 270 TAAAGCGGATTCAGCAAGCGGCTCCAGATAGTTGCGCTTTAGTTTCTGCTTGCAGCAAC 211
 QY 745 CATTTTCCGCGCTTGTGCAAGAGGACTTCAACCGCTTCCGCTTGTGCGCGGTTGGGT 804
 Db 210 CATTTTCCGCGCTTGTGCAAGAGGACTTCAACCGCTTCCGCTTGTGCGCGGTTGGGT 151
 QY 805 GCGCGATPACACACATCTGACCCAGTGTGATCTCGAGAGCTTTAAGCGGCTGACTCACGCG 864
 Db 150 GCGCGATPACACACATCTGACCCAGTGTGATCTCGAGAGCTTTAAGCGGCTGACTCACGCG 91
 QY 865 CGAGCGGGAATGGAAGGCTAAGAGCGCGCTTCAAGAGCTGCTTCAATCAATGATGA 924
 Db 90 CGAGCGGGAATGGAAGGCTAAGAGCGCGCTTCAAGAGCTGCTTCAATCAATGATGA 31
 QY 925 GAGCAAGTGTCCAGTTGAATGGGTTTCAAT 954
 Db 30 GAGCAAGTGTCCAGTTGAATGGGTTTCAAT 1
 RESULT 8
 BD165657/c
 LOCUS BD165657 870 bp DNA linear PAT 17-JAN-2003
 DEFINITION Novel polynucleotide.
 ACCESSION BD165657
 VERSION BD165657.1 GI:27871469
 KEYWORDS JP 2002191370-A/3456.
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 870)
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotide
 JOURNAL Patent: JP 2002191370-A 3456 09-JUL-2002;
 KYOWA HAKKO KOGYO CO LTD
 COMMENT OS Corynebacterium glutamicum
 PN JP 2002191370-A/3456
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
 KEIKO OCHIAI,
 PI HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
 OZAKI
 PI
 PC C12N15/09, C12N15/09, C07K16/12, C07K16/34, C07K16/40, C12M1/00, PC
 C12N1/15,
 PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/

PC 04.C12P13/08,
PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53, PC
G01N33/566,
PC G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),
PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),
PC C12N15/00,
PC C12N5/00,C12N15/00
CC Novel polynucleotide
PH Key Location/Qualifiers
FT source 1..870
FT /organism='Corynebacterium glutamicum'.
FEATURES
source
1..870
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 36.6%; Score 870; DB 6; Length 870;
Best Local Similarity 100.0%; Pred.No.9.5e-203; Indels 0; Gaps 0;
Matches 870; Conservative 0; Mismatches 0;
QY 85 AGCGCGCAATCCCTCGATTGCTGCAATCAAGCGCGCTCTGTAGTCTAGTAGATCT 144
DB 870 AGCGCGCAATCCCTCGATTGCTGCAATCAAGCGCGCTCTGTAGTCTAGTAGATCT 811
QY 145 AGATTCCAGCGGCATCGTTGCCAATACATCGGTGTCAATGGGTATCTATCGAGAG 204
DB 810 AGATTCCAGCGGCATCGTTGCCAATACATCGGTGTCAATGGGTATCTATCGAGAG 751
QY 205 GATCATCTCTCTGTTAGTATGGGAGCGAGCTTGGTTTCGGGAGAAAGTCCCCAAC 264
DB 750 GATCATCTCTCTGTTAGTATGGGAGCGAGCTTGGTTTCGGGAGAAAGTCCCCAAC 691
QY 265 AAGGCTCGGGGAATTCGCTACCAAAACCTTCGCGCGACGGGCAATGATAGCGCT 324
DB 690 AAGGCTCGGGGAATTCGCTACCAAAACCTTCGCGCGACGGGCAATGATAGCGCT 631
QY 325 GCGCCGCCACAGGACCATCGAGCGCGCTCAGGTACGGTCTTCAAGCAACCTTTGG 384
DB 630 GCGCCGCCACAGGACCATCGAGCGCGCTCAGGTACGGTCTTCAAGCAACCTTTGG 571
QY 385 ACCGAAAGGTAGAGCGGGATCCGACCCCATCTAGTTTCCCATCAACCATGTAGGATC 444
DB 570 ACCGAAAGGTAGAGCGGGATCCGACCCCATCTAGTTTCCCATCAACCATGTAGGATC 511
QY 445 CCGCATATGAGGGGTTGCAATGGCGCAAGTGGCGATGTTTCCAGTTCTTACTTACA 504
DB 510 CCGCATATGAGGGGTTGCAATGGCGCAAGTGGCGATGTTTCCAGTTCTTACTTACA 451
QY 505 TCCGCGCACGGGATTAGTTTCAAGGGTTACCGCTCTTAAACATCTCCACGCCGACAA 564
DB 450 TCCGCGCACGGGATTAGTTTCAAGGGTTACCGCTCTTAAACATCTCCACGCCGACAA 391
QY 565 GGATTAATGTGTGGCTTCATCTTCCAGCGCAGCGTGGAGGTGCTTCCACCCCAAGAGC 624
DB 390 GGATTAATGTGTGGCTTCATCTTCCAGCGCAGCGTGGAGGTGCTTCCACCCCAAGAGC 331
QY 625 TACCTCTGTAACACGGGAGGAACCATGTGGATAGCGAATCTGGGTTGATGGCGATGGT 684
DB 330 TACCTCTGTAACACGGGAGGAACCATGTGGATAGCGAATCTGGGTTGATGGCGATGGT 271
QY 685 TAACGGGATTTCAGAGAGCGTTCAGATAGTTTCGGCTTTAGTTTCTGCTTGCAGCAAC 744
DB 270 TAACGGGATTTCAGAGAGCGTTCAGATAGTTTCGGCTTTAGTTTCTGCTTGCAGCAAC 211
QY 745 CATTTTCGCGCTGCTTGCAGAGGACTTCACCGCTTCGGTTCGTTTGGCGGTGGGT 804
DB 210 CATTTTCGCGCTGCTTGCAGAGGACTTCACCGCTTCGGTTCGTTTGGCGGTGGGT 151
QY 805 GCGGATACCAACTCGACCCACGTCGTGATGTCGAGAGGTTTAAACGCGTGTACTACGCG 864
DB 150 GCGGATACCAACTCGACCCACGTCGTGATGTCGAGAGGTTTAAACGCGTGTACTACGCG 91

QY 865 CGAGGGGGAATTCGAAAGCGCTTAAGAGCGGCTTCGAGCTGCTTCATCATGATTGA 924
DB 90 CGAGGGGGAATTCGAAAGCGCTTAAGAGCGGCTTCGAGCTGCTTCATCATGATTGA 31
QY 925 GAGCAAGTGTCCAGTTGAATGGGGTTTCAT 954
DB 30 GAGCAAGTGTCCAGTTGAATGGGGTTTCAT 1
RESULT 9
AX063771 822 bp DNA linear PAT 24-JAN-2001
LOCUS
DEFINITION Sequence 53 from Patent WO0100843.
ACCESSION AX063771
VERSION AX063771.1 GI:12541483
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1
Pompejus,M., Kroeger,B., Schroeder,H., Zelder,O. and Haberhauer,G.
corynebacterium glutamicum genes encoding metabolic pathway
proteins
Patent: WO 0100843-A 53 04-JAN-2001;
JOURNAL BASF AKTIENGESSELLSCHAFT (DE)
FEATURES
source
1..822
Location/Qualifiers
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101..802
/note="unnamed protein product; RXA01394"
/codon_start=1
/transl_table=11
/protein_id="CAC25127.1"
/db_xref="GI:12541484"
/translation="MEIFITGLIGLLSISGPONVIVIKIGIKRELIALLVCLII
SVDFIFLIGLISNAAPIVLDIMRWGDIAYLLWFAVMAKDMNVEAPQIIE
ETEPVDDTPLGSAVATDTRVVRVEVSKORVWKPMLALVLTWLNPNAYLDA
FVPIGCAQYGDGTGRWIPFAAGFAASLWFLVGFCAALSRPLSPKVRWVNVV
AVMTALAKLMLNG"
ORIGIN
Query Match 34.6%; Score 822; DB 6; Length 822;
Best Local Similarity 100.0%; Pred.No.6.1e-191;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 925 GAGCAAGTGTCCAGTTGAATGGGGTTTCATGAGCTATATTAAACCATGTTAAGAACCA 984
DB 1 GAGCAAGTGTCCAGTTGAATGGGGTTTCATGAGCTATATTAAACCATGTTAAGAACCA 60
QY 985 TCATTTTACTTAAGTACTTCCATAGTTCAGATGGTATCATGAATCTTCATTACAGG 1044
DB 61 TCATTTTACTTAAGTACTTCCATAGTTCAGATGGTATCATGAATCTTCATTACAGG 120
QY 1045 TCTGCTTTTGGGGCCAGTCTTTTACTGTCTCATCGACCGCAGATGTTCTGTGATTAA 1104
DB 121 TCTGCTTTTGGGGCCAGTCTTTTACTGTCTCATCGACCGCAGATGTTCTGTGATTAA 180
QY 1105 ACAGGAATTAAGCGGAGGACTCATTTGGGTTCTTCTCGTGTGTTTAAATTTCTGAGT 1164
DB 181 ACAGGAATTAAGCGGAGGACTCATTTGGGTTCTTCTCGTGTGTTTAAATTTCTGAGT 240
QY 1165 CTTTTGTGTTTCAATCCCGCACCTTGGCGTTGATCTTTTGTTCATCGCCGCCGATCGT 1224
DB 241 CTTTTGTGTTTCAATCCCGCACCTTGGCGTTGATCTTTTGTTCATCGCCGCCGATCGT 300
QY 1225 GCTCGATATTATGCGCTGGGGTGGCATCGCTTACCTGTTATGTTTGGCTCATGGCAGC 1284
DB 301 GCTCGATATTATGCGCTGGGGTGGCATCGCTTACCTGTTATGTTTGGCTCATGGCAGC 360

1285 GAAAGACCCATGACAAACAAAGGTGGAAGCGCCACAGATCATTGAAGAAACAGAACCAAC 1344
 361 GAAAGACCCATGACAAACAAAGGTGGAAGCGCCACAGATCATTGAAGAAACAGAACCAAC 420
 1345 CGTGGCCGATGACACGCTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 1404
 421 CGTGGCCGATGACACGCTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCCACCGGT 480
 1405 GCGGTGGAGTGAGCGTGCATTAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCCAAT 1464
 481 GCGGTGGAGTGAGCGTGCATTAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCCAAT 540
 1465 CGTGTGACCTGCTGTAACCCGAAATGCTATTGAGCGCTTGTGTTATTCGCGCGGT 1524
 541 CGTGTGACCTGCTGTAACCCGAAATGCTATTGAGCGCTTGTGTTATTCGCGCGGT 600
 1525 CGCGCGCAATACCGCGACACCGACGCGGTGGATTTTGGCGGTGGCGGTTCGCGCAAG 1584
 601 CGCGCGCAATACCGCGACACCGACGCGGTGGATTTTGGCGGTGGCGGTTCGCGCAAG 660
 1585 CTTGATCTGGTTCCTCGCTGGTGGTTCGCGCGCAGCAGCATTTGTCACCCCGCTGTCAG 1644
 661 CTTGATCTGGTTCCTCGCTGGTGGTTCGCGCGCAGCAGCATTTGTCACCCCGCTGTCAG 720
 1645 CCCAAGGTGGCGCTGATCAACGCTGCTGCGAGCTGTCGAGATTCGAGACCGCATTTGCCAT 1704
 721 CCCAAGGTGGCGCTGATCAACGCTGCTGCGAGCTGTCGAGATTCGAGACCGCATTTGCCAT 780
 1705 CAACTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC 1746
 781 CAACTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC 822

RESULT 10
 AX244059
 LOCUS AX244059 822 bp DNA linear PAT 28-SEP-2001
 DEFINITION Sequence 51 from Patent WO0166573.
 ACCESSION AX244059
 VERSION AX244059.1 GI:15859123
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O., Haberhauer, G.,
 Kim, J. W., Lee, H. S. and Hwang, B. J.
 Corynebacterium glutamicum genes encoding metabolic pathway
 proteins
 JOURNAL Patent: WO 0166573-A 51 13-SEP-2001;
 BASF AKTENSELSCHAFT (DE)
 FEATURES Location/Qualifiers
 1..822
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101..802
 /note="unnamed protein product; RXA01394"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC8508.1"
 /db_xref="GI:15859124"
 /db_xref="REMBASE:CA8508"
 /translation="MEIFITGLLLGASLLISIQPNVIVKQIKREGLIALLVCLV
 SDVFLIAGTGLDLSNAAPVILNFWAVNAARDAMTKVPAQPIE
 EFTPTDPTPLGSSAVDTNRVRVSVSKORVWVKMLAIVLTLNPNAYLDA
 FVFTGGVGAQYDGTGRWIFPAAGAPASLIWFLVGVGAALSRPLSSPKWRWLVVV
 AVWTAIAIKLMNG"
 CDS
 ORIGIN

Query Match 34.6%; Score 822; DB 6; Length 822;
 Best Local Similarity 100.0%; Pred. No. 6.1e-191;

Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 925 GAGCAAGTGTCCAGTTCGAATGAGGGTTCATGAGCTATATTAACCAATGTTAAGAACCA 984
 Db 1 GAGCAAGTGTCCAGTTCGAATGAGGGTTCATGAGCTATATTAACCAATGTTAAGAACCA 60
 QY 985 TCATTTTACTTAAGTACATTCCTCATAGGTACAGATGGTGATCATGGAATTCCTTATACAGG 1044
 Db 61 TCATTTTACTTAAGTACATTCCTCATAGGTACAGATGGTGATCATGGAATTCCTTATACAGG 120
 QY 1045 TCTGCTTTTGGGGCCAGCTCTTTTACTGTCATCGACCGCAGAAATGACTGTGTGATTA 1104
 Db 121 TCTGCTTTTGGGGCCAGCTCTTTTACTGTCATCGACCGCAGAAATGACTGTGTGATTA 180
 QY 1105 ACAGGAATTAAGCGGGAAGACATTCGCGTTCCTTCGCTGTGTGTTAATTTCTGACGT 1164
 Db 181 ACAGGAATTAAGCGGGAAGACATTCGCGTTCCTTCGCTGTGTGTTAATTTCTGACGT 240
 QY 1165 CTTTGTGTTTCAATCGCGGACCTTTGGGCGTTGATCTTTTGTCCAAATGCGCGCGCATCGT 1224
 Db 241 CTTTGTGTTTCAATCGCGGACCTTTGGGCGTTGATCTTTTGTCCAAATGCGCGCGCATCGT 300
 QY 1225 GCTCGATTTTATCGCTGGGTTGACATCGCTTACCTGTTATGTTGTTTCCGTCATGCGACG 1284
 Db 301 GCTCGATTTTATCGCTGGGTTGACATCGCTTACCTGTTATGTTGTTTCCGTCATGCGACG 360
 QY 1285 GAAAGACCCATGACAAACAAAGGTGGAAGCGCCACAGATCATTGAAGAAACAGAACCAAC 1344
 Db 361 GAAAGACCCATGACAAACAAAGGTGGAAGCGCCACAGATCATTGAAGAAACAGAACCAAC 420
 QY 1345 CTTGCGCGATGACACGCTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCAACCGGT 1404
 Db 421 CTTGCGCGATGACACGCTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCAACCGGT 480
 QY 1405 GCGGTGGAGTGAGCGTGCATTAAGCAGCGGTTTGGTAAAGCCCATGTTGATGCCAAT 1464
 Db 481 GCGGTGGAGTGAGCGTGCATTAAGCAGCGGTTTGGTAAAGCCCATGTTGATGCCAAT 540
 QY 1465 CTTGCGCGATGACACGCTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCAACCGGT 1524
 Db 541 CTTGCGCGATGACACGCTTTGGCGGTTTGGCGGTGGCCACTGACACGCGCAACCGGT 600
 QY 1525 CGCGCGCAATACCGCGACACCGACGCGGTGGATTTTGGCGGTGGCGGTTCGCGCAAG 1584
 Db 601 CGCGCGCAATACCGCGACACCGACGCGGTGGATTTTGGCGGTGGCGGTTCGCGCAAG 660
 QY 1585 CTTGATCTGGTTCCTCGCTGGTGGTTCGCGCGCAGCAGCATTTGTCACCCCGCTGTCAG 1644
 Db 661 CTTGATCTGGTTCCTCGCTGGTGGTTCGCGCGCAGCAGCATTTGTCACCCCGCTGTCAG 720
 QY 1645 CCCAAGGTGGCGCTGATCAACGCTGCTGCGAGCTGTCGAGATTCGAGACCGCATTTGCCAT 1704
 Db 721 CCCAAGGTGGCGCTGATCAACGCTGCTGCGAGCTGTCGAGATTCGAGACCGCATTTGCCAT 780
 QY 1705 CAACTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC 1746
 Db 781 CAACTGATGTTGATGGTGTAGTTTTCGCGGGTTTGGAAATC 822

RESULT 11
 AP005218/c
 LOCUS AP005218 308650 bp DNA linear BCT 10-JUL-2003
 DEFINITION Corynebacterium efficiens YS-314 DNA, complete genome, section 5/11.
 ACCESSION AP005218 BA000035
 VERSION AP005218.1 GI:23493016
 KEYWORDS
 SOURCE Corynebacterium efficiens YS-314
 ORGANISM Corynebacterium efficiens YS-314
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 Nishio, Y., Nakamura, Y., Kawarabayashi, Y., Usuda, Y., Kimura, E.,
 1
 REFERENCE
 AUTHORS

Thu Mar 18 12:30:58 2004

Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and
 Gojobori, T.
 Comparative Complete Genome Sequence Analysis of the Amino Acid
 Replacements Responsible for the Thermostability of *Corynebacterium*
efficiens
 Genome Res. 13 (7), 1572-1579 (2003)
 22723752
 12840036
 2 (bases 1 to 308650)
 Kawarabayashi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and
 Director-General of Biotechnology Center.
 Submitted (17-MAY-2002) Director-General of Biotechnology Center,
 National Institute of Technology and Evaluation, Biotechnology
 Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan
 [E-mail: bio@nitech.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424]
 Kawarabayashi, Y. is officially affiliated with the National
 Institute of Advanced Industrial Science and Technology, Tsukuba,
 Ibaraki, 305-8566 Japan
 Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the
 National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
 Itoh, T. is at the Japan Biological Information Research Center,
 Koto-ku, Tokyo, 135-0064 Japan
 Yamagishi, A. is at Tokyo University of Pharmacy and Life Science,
 Hachioji, Tokyo, 192-0392 Japan
 Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co.,
 Inc., Kawasaki, Kanagawa, 210-8681 Japan
 The other authors are at the National Institute of Technology and
 Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.
 Location/Qualifiers
 1..308650
 /organism="Corynebacterium efficiens YS-314"
 /mol_type="genomic DNA"
 /strain="YS-314"
 /db_xref="taxon:196164"
 complement(83)..241)
 /note="CE1178"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC17989.1"
 /db_xref="GI:23493017"
 /translation="MTTGIILGTGPGDAGLKGISDIPRGVGHFRFRKPFVALQYR
 TGAGAND"
 365..742
 /note="CE1179"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC17989.1"
 /db_xref="GI:23493018"
 /translation="MEKSQLQKQKQDFASVDHLSDDAAMFIDNELSRGAMHRAHLH
 LVHCACRQELARQDVTDFVTECNQNVTPSDLKAKMLSLANECAPGAEADAI
 QRPESFVAKVSVAVRRTQGR"
 774..1271
 /note="CE1180, similar to AJ292256-1|CAC21724.1| percent
 identity: 26 in 155 aa"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC17990.1"
 /db_xref="GI:23493019"
 /translation="MFSNVGGEILLFTVGLIVIGPERLPRMLDLKAAIYAARTAI
 NNKQNLDTDFGEEDFELRKPMQIAQIRSMSPKTAITRALFDGDSFLDSFDPKKIM
 ADDTAGAQRHQVAKTGTTPPEAEVVERPADRIDTQSPWTPGKNDGKPKAGGGYT
 WTDLI"
 complement(1353)..2480)
 /note="CE1181, similar to AL390975-15|CAC01353.1| percent
 identity: 63 in 370 aa"
 /codon_start=1
 /transl_table=11
 /product="putative ATP-binding protein Mrp"
 /protein_id="BAC17991.1"

/db_xref="GI:23493020"
 /translation="MSQVTEASVRSALSRVEDPEIGKIPITELGMVKSVAIDGSDVQV
 EVLITIAACPMKTTIVNTAAIKDIEGVGDVQVTTDMSDEQRRLRSLRGNTADP
 VIPPAQPTTTRVAVASGKGVSGKSMFTVNATAPAAKGLSGVILDDADIYGHVSFPCW
 LGSADPHQVDMVMPPOAHGVMSIAHFVEGNAVPPVWRGPMHRAICQFSLDFWFG
 DLDTLLDLPCTGDIATIAQLLPNAELIIVTTPQAAADVAERAGSLSLQTKQKVA
 GVENMSAMVDPDSTMEVFGSGGKIAQRLSMLTGEDVPLVLSVPLDPSLRVGGDV
 GNPIASEPSPPTTAINATADTLAKRKYSIVGKTLGLGVK"
 complement(2513)..3271)
 /note="CE1182"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC17992.1"
 /db_xref="GI:23493021"
 /translation="MTTATPWADAFPHVGDLPQLLEVDHTERAVAHLOEYFATDPT
 GPMYTGSLFRNLRGGGDAQPVNRNITKDDLSLLDVPFGVQLLEPPEDAQSI
 DLEEVTRDYAVELPQIEQIRGLASPTDVLGASKAAAEHLDKGNLLWTAPENG
 FGPRTRSKLTARKRPKLPFIDSAIKRQNTNSVGFEYFENFTVLRAEDKLSHLKGI
 RTEADSGNSALADLSITRVFDVVVVMADHKRT"
 complement(3282)..3836)
 /note="CE1183, similar to AE007003-3|AAK45527.1| percent
 identity: 57 in 159 aa"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC17993.1"
 /db_xref="GI:23493022"
 /translation="MADYSRSLDPLVGTGRKRFKEDDVTGAVAKVAFPGTQY
 LFWQITFVIMINTGFFANWDYPFILLNKFSTQAAVAAPIILLANRQEDRK
 STILADRRABETKADTEFTRELAVGEMVTRDYLRHLDLDRGLERLEAKL
 DDAVASRNDPDLNAPTOGKVD"
 complement(3829)..5124)
 /note="CE1184, similar to AE007003-4|AAK45528.1| percent
 identity: 52 in 417 aa"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC17994.1"
 /db_xref="GI:23493023"
 /translation="MRGITLTVAGLACMVVRGDDTVIGRVDDVVVTRSGQDSRV
 LGLVVELVNNKRIFLPMRLIAAIEKVENLISVLSASFKARFGELOVIGDIOTK
 VHTDPELQLHGRVAVIEADVELEQTRTDWVSRVALIGERKFGREPTLYTAPWSH
 VHGTAGGAGETNQTAEIAGFEDMAADAKELYGLPKQPVVAGVLEDERLADIL
 QELSEDRQAELETLTIERAADVLEMDPDAADLGLPEKDGAGVLELMDPEPSAP
 VRLMGFNEDTVGALMTPELTWDPSTTVABALAMARNPDLTSLASLVFVREPTAT
 PTGKYLGCVHLQGLRPPESVIGLIDPLPLVANDSOETAAARFFATYVNLVCGPYL
 DENGHLIGAVAVDILLDLMLPEDWDRDAGRPGVKEEILG"
 5264..5782
 /note="CE1185, similar to AE007003-6|AAK45530.1| percent
 identity: 60 in 155 aa"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC17995.1"
 /db_xref="GI:23493024"
 /translation="MSMOATGSGNNRAELIGLKKRPGCVPGSFQYEEAQRVDMLS
 DNEFFVAELITVGLMEVETVTRGLTWKRVLAGGAAGMLGIFFLGVLGVSNLWP
 SSLITGLVMGVPGVLAAPVYWASRKRDFTSATQIVAGRYDILCRDPRAREADMI
 ARNMGEIGHPRQ"
 5805..6500
 /note="CE1186"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC17996.1"
 /db_xref="GI:23493025"
 /translation="MRTGVKALLGLVAVLGGLLVAGSVDPNRPPEHPQVAEQK
 ENLYTAVDPDYLGHNVISLYKTVNGRFAVIRHEDNLADFPDNEGTSDLVIG
 CTGALLEVYVNPVAGLEBEELYAAVEAGETDINSGEWRDFTDALIGSPSLRDTADP
 SNAQCGPGANGLFQNVPIYRNPIFSRADREILNWSAGIATTSLOELIEAEKELRS
 RELVKEFVASKGG"
 6660..7373

	/note="CE1187, similar to AE007262-3 AAK65408.1 percent identity: 32 in 235 aa" /codon_start=1 /transl_table=1 /product="conserved hypothetical protein" /protein_id="BAC17997.1" /db_xref="GI:23493026" /translation="METILIGVLLFVAACVOGAGFGELGLAAPLIALMPDLPGLI LLARGLSLATNABRGREVPYHVLISLRISULGVWAVAFVAGIKVLATAV ILGLTILSVGSPPGRGNRSFLAGSGPLGTITRIGCPPLALIMRSMSPRVRLTS VTIFLSIGSAILLTGAGALGMHLQALVALPAPAVIAGFYLGSVWNKLNRLLFI LGS 7592.. 8149
CDS	/note="CE1188" /codon_start=1 /transl_table=1 /product="hypothetical protein" /protein_id="BAC17998.1" /db_xref="GI:23493027" /translation="MGRGDGPAREPRPADRCGRKTIKWKTKGAIRAAARALLLV GGGFTMSWDEVTGSDORIAAGNLASRAQKWTINGQDVDDNNVAKVPEPTLIV FQQLISTETSGDNLSALTLPFGSTTGATADARAADALNELMDITTDUGLELPSITG SNTYTTEGGVRGSGVSALSIIKQAA" complement(8190..9443) /note="CE1189, similar to AF024666-23 AAG03374.1 percent identity: 84 in 408 aa" /codon_start=1 /transl_table=1 /product="transposase" /protein_id="BAC17999.1" /db_xref="GI:23493028" /translation="MAQGPSIDPTYILLDELAQSPDLRMENLOGFINQILSQADQ VCGDIATFVQDKRWNRTRNGVDIPRVGITDVAHPKRTGSGTFDFDWMLLERTRAER ALCTTATYLKGSVTRNDRIKVNINLIKSQSVDMAKDOLDMVEEFPRTPLDR PLYLYSCDALTMVKREGRRVHTSVLLATGVNAEGRELLGMQNATAESVASMTGFR DLKARKDELTVTLGHQIAEVLHPNASWRQRTHFSKNL YGVMSKTOWTSL AMPHFITQQDAOSVWGQAEVFAFCQRFEDVADYLEALDELLAFNPASKVMYKV WSNNTPIQRRELRRTRDVDGIFPNRDADVRLGVAVLAEQHDDMTIQXRYMSLTSL EQ TRAMWTANIDSDDITSRIIRKOVA" complement(11439..15143) /gene="odha" CDS complement(11439..15143)
	Query Match 34.5%; Score 818.6; DB 1; Length 308650; Best Local Similarity 68.5%; Pred. No. 7.Re-190; Matches 1182; Conservative 0; Mismatches 524; Indels 19; Gaps 3;
QY	1 CCATTTCCTGAAGTGTACTTCCTGCCTGGCCCAATTCCTTCGGCGGCGAAGAAGTGA AAAAC 60
DB	192625 CCTTCGCGAGGGTGTCA CGCTGC CGCGGCCGATACCTCCGCTGGAAGGTCAAGAAC 192566
QY	61 CTTGACCTTTTCAGAGTAAC TAAGCGCGCAATCCCTCGATTCGTGCATCAAGCGGC 120
DB	192565 CCGAACCGTTCAGGTTAGTTAGTT---CTCAGACC CGCGCGCATCACACGGC 192510
QY	121 GTCTGTAGTCTAGCTAGAGATCTAGATCCAGCGCCCATCTGTGGCCAATACATCGGTGT 180
DB	192509 GTCTGTAGTCTGTGGCGAGCGCCTCGACTCCAGCGCCCATCTGTGGCCAATACAGCGGGGT 192450
QY	181 GTCAA TGGGTATCTCATCGAGGAGGATCACTTCTCTGCTTTTAGCATGGAGCAGCTTG 240
DB	192449 GTCCACACCTTCTCATCGCTGCACCAACATCCCGCGAGCAGCATCGCGCGCGCTG 192390
QY	241 GGTTCGCGNAGAGTCCCACCAACGAGCGCTCGCGCAATTCGCTTCCACAAAACCTCCGC 300
DB	192389 GGCTTCGGGACAGAGGCCCAACCCCMAGCGGAGTTCGACCGCTTCCCCGHAACCTTCGC 192330
QY	301 CGACGGACAATGGATAGCGGCTGCGCCCCACAGGACCATCGACGCGCGCTCCAGGTC 360
DB	192329 AGACGGCACACCGAGACAGCGGCGCTCGCGAGCGGCTCCGCTCGACCTCCCTCGCATC 192270
QY	361 ACGGTCTTGAAGCAATCTTTTGGGACCGAGCGGTPAAGACGGGCAATCGAGGCCCAATCTAG 420
Db	192269 GCGGTCCTGCGAGCAATCAATTCGSGGCCGAAACGCAATACCGGCAATCCCGCATCGGG 192211
Qy	421 TTTCCTCATCAACCAATGTAGGCAATCCCGCAATAGAGGGGGTGTSCAATGCGCAAGTGGCGCAT 480
Db	192209 TTGGCCATCCACGGTGTATCCGTGCGCTCAGTTCGSGGGGTGGCCACCGGCAATGTCTCAT 192150
Qy	481 GGTTCACAGTGTACTACTTTCACATCCCGCCACGAGATAGCTTTCACGGGTTCACGGCTCC 540
Db	192149 GACCCGAGGGGTAGGACCTCACACCCGGCAACGGGGTCCGCTCCGGGGTCAACGGGCC 192090
Qy	541 TAAACAATCTCAGCGCGAGCAGAGATATATGTGTGCGCTTCTATCTTCCAGCCCGACGT 600
Db	192089 GAGCAGGAAACCGCGGCGCAACAGGGAACGGGTGTGGGGCTTCGTCTCCACACGACGCGT 192030
Qy	601 GAGCGTGTCTCCACCCCAAGAAAGCTTACTCTGTGTGAACAACGAGGAGAAACATGTGGATAG 660
Db	192029 CAGGCTGACCCGACCCCAATGTGCAACCTCGCGCAACACGGGCGGGAACAGAGTGGACAG 191970
Qy	661 CGAATCTGCTGTANTGCGCATGTTTAAACGGGATTTTACGAGAGCGTTCAGATATGTTCGGC 720
Db	191969 GBAATCGGGTGTATAGCCACCGTTCAGGGGATCTCGTCAAGCGGTTCGCGCATGTCTC 191910
Qy	721 TTTAGTTTCTGCTTCGAGCAACAACATTTTCGCGCTGCTTTCGACAGACTTCACCCGC 780
Db	191909 ACGGCTCTCCGCTTCGAGCAGCGCATTTTGGGGCGCGCTTCGACGACGACTCCCCGCG 191850
Qy	781 TTCGCTGTCTTTGGCGGGTGGGTGCGCATACCAACACATCGACCCACATGATCTCTCGAG 840
Db	191849 TTCGTTGGCAACCGCGGCTGGGTGCGCGACACCACTAGTACCCGACCGAGGATTTTTCGAG 191790
Qy	841 AGCTTTAAACGCGCTGACTCACCGCGAGGGGGGAATGAAAGGSGTCAAGGAGCGCCTTC 900
Db	191789 TGCTTTGATGCTGTGCTGACCCGCGAGGGGGAGATCGACAGTTCGACGGGAGGGGTCTC 191730
Qy	901 GAAGTCGCTTCATCAATGATTTGAGAGCAAAAGTTCACAGTTGTAATGGGGTTCATGAAGCT 960
Db	191729 GAAGCTGCGCTCTCGCATGATGGTCAGGAGGGTTCACAGTGGATCGGGTTCATGAAGCA 191670
Qy	961 ATATTAAACCAATGTTAAGAACCAATCAATTTTACTTAACTACTTCCATAGTACACGATGGT 1020
Db	191669 ATTCTAAACCAATGTTCACTCTCATCAATTTTACTTAATGTGTTGTGCGGCAATGCG -191610
Qy	1021 GATCATGGAATCTTCATTAAGTCTGCTTTTGGGGCGCAGTCTTTTACTGTCCATCGG 1080
Db	191609 GBACTGGAAATCTTTGTGTCACCGTTTTGTTGTGGAGCCAGTCTGCTGTGGCGCATCGG 191550
Qy	1081 ACCGCAAGATGACTGTGATTAACAAGGAATTAAGCGCGAAGGACTCATTCGGGTCT 1140
Db	191549 CCCACAGAATGTCTGTGTATCAACAGGGCATCAACACGAGGGGATCAACGGCGTTCAT 191490
Qy	1141 TCTGTGTGTTTAAATTTCTGAGTCTTTTGTTCATTCGCGCAGCACTTCGGGCTGTGATCT 1200
Db	191489 CATCGTCTGTCTGCTCCGAGGTGGTCCCTGTTTCCCTTCGCGACCTTCGGGTCGGCT 191430
Qy	1201 TTTGTCAATCGCGCCCATCTGCTCGATATATGTCGTGGGGTGGCGGTGCGATCTTACCT 1260
Db	191429 GATCTCCGACACCGCCCGCATCAITTTCTCGACATCTCTGCGCTGGGTGGGTCTGCTACCT 191370
Qy	1261 GTTATGTTTTCGCTCATGCGCAGCGAAAGACGCGCATGACAAACAGGTTCGAAGCGCCACA 1320
Db	191369 GCTGTGTTTCGCGTGTATGGCGCGCGCGACGCGCTTCGCGCCCGCACCGAGGTAAACCTT 191310
Qy	1321 GATCATTTGAAGAAACAGAACCAACCGTGCCGATGACACGCTTTTGGCGGTTTCGGCGGT 1380
Db	191309 TGTG-----GAGGATTCCGAACCCGTG---CGGACCGTTCGCTCCGGCGG 191265
Qy	1381 GGCACTGACACGCGCAACCGGGTGGGGTGGAGGTGAGCTCGATAAGCAGCGGGTTTG 1440
Db	191264 GGGCGTGAACGACGAAAAACACGACCCCGCGCTCCGCAATCATATAGGACCCCGCAGGCTG 191205
Qy	1441 GGTAAAGCCCAATGTTGATGGCAATCGTGTGACCTGTGGTGAACCCGAATGCTATTGGA 1500
Db	191204 GGTCAAGCCCAATGCTCATGCGCAATGTGCTGACCTGCTGCTCAATCCCAATGCTCTACCTGGA 191145

ORIGIN			
Query Match 33.6%; Score 798.8; DB 1; Length 1771;			
Best Local Similarity 68.4%; Prid. No. 3.3e-185;			
Matches 1158; Conservative 0; Mismatches 517; Indels 19; Gaps 3;			
QY	32	CAATTCCTCGCGGCGCAAGAGTCGAAAGACCCCTGAAACCTTTTCAGAAAGTAACTAAGCGCGC	91
DB	1	CGATACCTCCGCTGAGAGGTCAAGAACCCGGAACCGTTCCAG---GTTTAGGTTCTC	56
QY	92	AATCCCTCGATTGCTGCATCAACGACGGCGTCTGTGAGTCTAGCTAGAGATCTTAGATGCC	151
DB	57	AGACCCGCGCGCGCGCATCCACACGGCGGTCTGTGAGTCTGGCGAGCAGCTTGGAATCC	116
QY	152	AGGGCGCATCGTTGGCCAAATCATCGGTGTGTCAATGGGTATCTCATCGAGAGGATCACT	211
DB	117	AGCGCCCATGTTGTCANATACAGGGGGGTGCCACCACTTCTCATCGAGTGCACCA	176
QY	212	TCTCTCGCTTTTAGCATCGGACGAGTCTGGGTTCGGGAAAGAGTCCCCACACAGGCGCT	271
DB	177	TCCCCGCGCAGCATCGATCGCGCGCTCGGCGCTCGGCGCAGAGCGCCACCCCGCGCG	236
QY	272	CGGCGAATTGCTCTACCAAAACCTTCGCGCGACGGGCAATGATACCGCGCTGCGCCCG	331
DB	237	AGTGCACACGCGCTCCCGGAACCTTCGACAGCGGCTCACCGAGACGCGCGCTGGCG	296
QY	332	ACAGACCATCATGACGCGCGCCGCTCCAGGTCAACGCTTTGAAAGCAATCTTTGGGACCGAG	391
DB	297	ACGGCTCGTCTGACCGCTGCGCTCGAGATCGCGTCTCTGACGACATCATTTGGGCGCGAA	356
QY	392	CGTTAGACGGGATCGGACGCCAATCTAGTTTCCCATCAACCATGTAGGCATCCCGCAAT	451
DB	357	CGCAGTACCGGCGATGCGACCCAGTGGGTGTGGCCCTCCACGGTGTACGGTGCCTCAGT	416
QY	452	GAGGGGGTTGCAATGAGCCAAAGTGGCGCATGGTTCCAAAGTTCTACTATCAATCCCGCG	511
DB	417	TCCGGGGTGGCACCGGCGAGGTGTCTCATGACCCCGAGCGGTAGGACCTCACACCGCGCA	476
QY	512	ACGGGATTAGCTTACCGGGTTACCGCTTCTAAACATCTCACCGCGAGCAAGATTAAT	571
DB	477	ACGGGGTCCGCTCGCGGGTCAACGCGCCCGAGACGGAAACCGGGCGACACAGGGAACGG	536
QY	572	GTGTGCGCTTCATCTTCCAAAGCGGACGTGAGCGTGTGCTCCACCCCAAGAAAGTACTCG	631
DB	537	GTGTGGGCTCGTCTTCCACGACGAGCTGAGGTGACCGCACCCCAATGTGCAACCTCG	596
QY	632	TTGAACAACGGAGGAAACATGTGTGATAGGAAATCTGGGTGTATGCGATGTTAAACGGG	691
DB	597	GCGAACACGGCGGGGAACACAGTGGACGGGAATCGGGGTTTGATAGCCACGTCAGGGGG	656
QY	692	ATTTCAGACAGGCGTCCAGATGTTGGCTTTAGTTTCTGCTTGCAGCAACACCATTTTC	751
DB	657	ATCTGTCAACGCGTTCCCGCAGTTGTCTACGGGTCTCGGCTCTGAGCAGCGCATCTTG	716
QY	752	CGGCTGTCTTCCACAAAGCATTCACCGCGTTTGGTGTGTTGGCGGTTGGTGGCGGAT	811
DB	717	CGGGCGGCTGACACGACCACTCCCGGCTTCGGTGGCAACCGCGCGGCTGGGTGCGCGAC	776
QY	812	ACCAACATCGACGCCGTGTGCTCGAGAGCTTTAAACGGCGCTGACTCACCGCGCGAGGGG	871
DB	777	ACAGTACCCACCGACGGAATTTCCACGTGCTTGTATGTCGTGTGCTGACCGCGCGAGGG	836
QY	872	GAAATGGAAAGGGCTAAGGAGCGCGCTTCGAAGCTGCTCTCATCAATGATTGAGAGCAAA	931
DB	837	GAGATCGACNGTGGAGGGAGGCGCTTTCGAAGCTGCCCTGTCGATGATGTGTGAGGAGG	896
QY	932	GTGTCCAGTTGATGGGGTTCATGAGCTATATAAACAATGTTAGAACCAATCATTTT	991

QY	1501	CCGCTTGTGTTTATCGCGCGCGTTCGGCGCAATACGGCACACCGGACGGTGGATTTT	1560
Ddb	191144	TGCCCTTGTCTTTCATCGGFGGTGTCGAGGCCAGTACGGGAGACCGTCTGGTGGATCTT	191085
QY	1561	CCGCGCTGGCGCGCTTCGGCGCAACCTGATCTGGTTCGCCGTCTGGTGGTTTCGGCGCAGC	1620
Ddb	191084	CGCTCGGGTGCCTTCGCGCGCAGCTGCTGCTGGTTCCTCCACTGGTCTGACGGCGGC	191025
QY	1621	AGCAPTGTACGCCCGCTGTTCAGCCCAAGGTGTGGCGCTGGATCAACGTCTGCTGGC	1680
Ddb	191024	CGCACTGTGCGTCCCTCTTCAGCGCGCGGTCTGGCGCTGGATCAACATAGTGTGGC	190965
QY	1681	AGTTGTGATGACCGGATTTGGCCATCAACTGATGTGTGATGGTTA	1725
Ddb	190964	CGTGTGTCTACCGGATTGGCGGTGAAGCTGATCCTGATGGTTA	190920
RESULT 12			
LOCUS	AB083133	1771 bp	DNA linear BCT 06-APR-2002
DEFINITION	Corynebacterium efficiens lysG, lysE genes for lysine export transcriptional regulatory protein, lysine exporter protein, complete cds.		
ACCESSION	AB083133		
VERSION	AB083133.1	GI:20065738	
KEYWORDS	Corynebacterium efficiens		
SOURCE	Corynebacterium efficiens		
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
REFERENCE	1. Itaya, H., Kimura, E., Kawahara, Y. and Sugimoto, S. lysG, lysE of Corynebacterium efficiens Published Only in Database (2002)		
AUTHORS	2 (bases 1 to 1771)		
JOURNAL	Itaya, H., Kimura, E., Kawahara, Y. and Sugimoto, S. Direct Submission		
REFERENCE	Submitted (04-APR-2002) Hiroshi Itaya, AJINOMOTO CO., INC, Fermentation & Biotechnology Laboratories, 1-1, Suzuki-Cho, Kawasaki-Ku, Kawasaki, Kanagawa 210-8681, Japan (E-mail:hiroshi.itaya@ajinomoto.com, Tel:81-44-244-7123 (ex.4146), Fax:81-44-222-0129)		
AUTHORS	Location/Qualifiers		
JOURNAL	1. .1771		
FEATURES	source		
	/organism="Corynebacterium efficiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:152794"		
	complement(47..919)		
	/gene="lysG"		
	complement(47..919)		
	/gene="lysE"		
	/codon_start=1		
	/transl_table=11		
gene	/product="Lysine export transcriptional regulatory protein"		
	/protein_id="BAB88826.1"		
	/db_xref="GI:20065739"		
	/translations="MNPHTDITLTIIDEGSFENASLALSIPSAVSQRILEKLEKSVG RYLPRTQPAVATGTEVQAKKMLQAETREQLAERLIDETPLTVAINDLSLST FPPVPEVAHWGAVTITLRFDEAHTLILRRGSLVGTREADPVAGCEVIRLGVWR HLPAVATFELPARYVDQDWRVPLRPGNDVLQDELDGKVDCAVARRRVSFPS AEGFGEAVRLGLGGLLPEAQFMLAAGDVVLQDEKVTDTPLTQWQWLESRLALR TDAAVDAARAGLTR"		
	981..1676		
	/gene="lysE"		
	981..1676		
	/gene="lysE"		
	/codon_start=1		
	/transl_table=11		
CDS	/product="Lysine exporter protein"		
	/protein_id="BAB88827.1"		
	/db_xref="GI:20065740"		
	981..1676		
	/gene="lysE"		
	981..1676		
	/gene="lysE"		
	/codon_start=1		
	/transl_table=11		
	/product="Lysine exporter protein"		

Db 897 GTCTCCAGTGGATCGGGTTTCATGAAGCAATCTTAAACATGTTCTAGTCTCTATCATTTT 956
 Qy 992 ACTTAAGTACTTCCATAGGTTCAGATGATCATGAAATCTTCTATACAGGTCTGCTT 1051
 Db 957 ACTTAAGTGGTTTGTGGGACATCGGACATGAAATCTTGTACCGGTTGTTG 1016
 Qy 1052 TTGGGGCCAGTCTTTTACTGTCTCATCGGACCGGACATGATGATGATTAACAGGA 1111
 Db 1017 TTGGGAGCCAGTCTGCTGTGGCCATCGGCCACAGATGCTCTGGTGTATCAACAGGC 1076
 Qy 1112 ATTAAGCGGAGGACTCATGCGGTCTTCTGCTGTGTAAATTTCTGACGTCTTTTG 1171
 Db 1077 ATCAAGCGGAGGACATCAGCGCTCATCATCTGCTGTCTGCTCCGAGTGTGCTTG 1136
 Qy 1172 TTCTATGCGGACATCTGGGCTTGTATCTTTTCTCAATGCGGCGGATGCTGCTGAT 1231
 Db 1137 TTACCTCTCGGACCTCTGGGCTGCGGCTGATCTCGACACCGCCGATCATCTTCGAC 1196
 Qy 1232 ATTATGCGTGGGTTGCGATCGCTTACCTGTATGTTTGGCTTCATGCGACGGAAGAC 1291
 Db 1197 ATCTGCGCTGTGCGGATCGCTTACCTGTGTGTGCTGCGGCTGATGCGGCGGCGAC 1256
 Qy 1292 GCGATGACAAACAGGTGGAAGCGCCACAGATCATTTGAAGAAACAGAACCAACGTCGCC 1351
 Db 1257 GCGTGGGCGCGGACCGGAGTAACTTTGTC-----GAGCATTCGACCC 1304
 Qy 1352 GATGACACGCTTTTGGCGGTTTGGCGGTTGGCCACTGACAGCGCAACCGGGTGGGGTG 1411
 Db 1305 GTTG---CGGACAGGCTCGGCTCGGCGGCGGCGGTGACAGCAAAACACCGCCGCTC 1361
 Qy 1412 GAGTGAGCTGATAGACGAGCGGTTTGGGTAAGCCCATGTTGATGCAATCGTCTG 1471
 Db 1362 CGCATCACATCAGGACCGCGGAGGTCTGGGTGAGGCCCATGCTCATGGCATTTGCTG 1421
 Qy 1472 ACTGTGTGACCGAATGCGTATTTGAGCGGTTTGTGATGCGGCGGCGTGGCGCG 1531
 Db 1422 ACTGTGCTCAATCCATGCTACCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1481
 Qy 1532 CAATAGCGGACACGCGGCGGTTTTCGCGGCTGCGGCGGTTTCGCGGCAAGCTGATC 1591
 Db 1482 CAGTACGGGAGACCGGTCTGGTATCTTCGCTGCGGCTGCTTCGCGCGAGCTGGTC 1541
 Qy 1592 TGGTTCGCGCTGGTGGTTTCGCGGACGATGTCACGCGGCTGTCGAGCGCCCAAG 1651
 Db 1542 TGGTTCGCGCTGGTGGTTTCGCGGACGCGGCTGTCGCGGCTGTCGAGCGCGCG 1601
 Qy 1652 GTGTGCGCTGGATCAAGCTGCTGCTGCGGTTTGTGATGACCGATGCGGCTCAAACTG 1711
 Db 1602 GTCTGCGCTGGATCAAGCTGCTGCTGCGGTTTGTGATGACCGATGCGGCTGAAGCTG 1661
 Qy 1712 ATGTTGATGGTTA 1725
 Db 1662 ATCTGATGGTTA 1675

RESULT 13
 LOCUS BX248357/c 349535 bp DNA linear BCT 06-NOV-2003
 DEFINITION Corynebacterium diphtheriae gravis NCTC13129, complete genome;
 segment 4/8
 ACCESSION BX248357 BX248353
 VERSION BX248357.1 GI:38199912
 KEYWORDS complete genome.
 SOURCE Corynebacterium diphtheriae
 ORGANISM Corynebacterium diphtheriae
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1 (bases 1 to 349535)
 AUTHORS Cerdeno-Tarraga, A.M., Efratratou, A., Dover, J.L.G., Holden, M.T.G.,
 Pallen, M., Bentley, S.D., Beara, G.S., Churcher, C., James, K.D., De
 Zoysa, A., Chillingworth, T., Cronin, A., Dowd, L., Feltwell, T.,
 Hamlin, N., Holroyd, S., Jagels, K., Moule, S., Quail, M.A.,
 Rabinowitch, E., Rutherford, K., Thomson, N.R., Unwin, L.,

TITLE Whitehead, S. and Barrell, B.G. Parkhill, J.
 The complete genome sequence and analysis of Corynebacterium
 diphtheriae NCTC13129
 JOURNAL Nucleic Acids Res. 31 (22), 6516-6523 (2003)
 PUBMED 14602910
 REFERENCE 2 (bases 1 to 349535)
 Cerdeno-Tarraga, A.M.
 Direct Submission
 TITLE Submitted (03-OCT-2003) Cerdeno-Tarraga A.M., submitted on behalf
 of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust
 JOURNAL Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:
 amctesanger.ac.uk
 FEATURES
 source 1. 349535
 /organism="Corynebacterium diphtheriae"
 /mol_type="genomic DNA"
 /strain="NCTC13129"
 /db_xref="taxon:1717"
 /note="biotype gravis"
 76..402
 /locus_tag="DIP1058"
 76..402
 /locus_tag="DIP1058"
 /note="Similar to Mycobacterium tuberculosis hypothetical
 11.0 kDa protein Rv1898 or MT1949 or MTCY180.20c
 SW:1192_MTCY180 (07734) (102 aa) fasta scores: E():
 5.8e-16, 56.43% id in 101 aa, and to Aquifex aeolicus
 hypothetical 17.9 kDa protein AQ_2067 TR:067847
 (EMBL:AE000771) (157 aa) fasta scores: E(): 4.7e-08,
 40.19% id in 102 aa"
 /codon_start=1
 /transl_table=11
 /product="Conserved hypothetical protein"
 /protein_id="C9E49578.1"
 /db_xref="GI:38199913"
 /translation="MLIDMLLAFSAVPTETPNNAEMADVAEALRVRESGLPNETN
 AMFLIIGEDVEVNAVKKATDAVLAASPRASLVKADIRPGTIGQLQKVASVERVL
 AGDSSES"
 88..372
 /locus_tag="DIP1058"
 /note="HMPfam hit to PF01910, Protein of unknown function
 DUF77"
 complement (405..4762)
 /note="Putative iron uptake system. Low GC content
 (58.02%)"
 complement (405..1226)
 /locus_tag="DIP1059"
 complement (405..1226)
 /locus_tag="DIP1059"
 /note="Similar to Streptomyces coelicolor putative
 iron-siderophore uptake system ATP-binding component
 SC151_25C TR:Q98215 (EMBL:AL109848) (301 aa) fasta scores:
 E(): 1.9e-42, 50.58% id in 255 aa, and to Escherichia coli
 ferric enterobactin transport ATP-binding protein fepC
 B0588 SW:FEPC_ECOLI (P23878) (271 aa) fasta scores: E():
 1.1e-41, 49.61% id in 262 aa"
 /codon_start=1
 /transl_table=11
 /product="Putative iron-siderophore uptake system
 ATP-binding component"
 /protein_id="C9E49580.1"
 /db_xref="GI:38199914"
 /translation="WISTSPILKTSVTLANDKHTSTLTVEVQSQQTALIGPNC
 GKSTLLKSFARILQHTQVTLGNPDLSTLTQIARALLPQTLAPADIVIELV
 GSGFPVYTLRWSTKDAQVDSALLATGAQLRRHVRTELSSGQVRVLANLAQ
 NTPVLLDEPTTLDVAHQQLLELALRQQLGRITVTLVLDLQCARVADHLVIMK
 GCAVYAGCAPSSIMTPQLISEVFDINVDHTIDHVVIVKTPMPQDPVSLNLQ"
 complement (564..1133)
 /locus_tag="DIP1059"
 /note="HMMsmart hit to SM00382, ATPases associated with a
 variety of cellular activities"
 complement (567..1130)
 /locus_tag="DIP1059"
 misc_feature
 misc_feature

misc_feature	/note="HMPfam hit to PF00005, ABC transporter"	
	complement(579..797)	
misc_feature	/locus tag="DIP1059"	
	/note="ProfileScan hit to PS0100, 2nd half motif for nucleotide binding, associated with P-loop."	
misc_feature	complement(753..797)	
	/locus tag="DIP1059"	
misc_feature	/note="ScanRegexp hit to PS00211, ABC transporters family signature."	
	complement(1047..1124)	
misc_feature	/locus tag="DIP1059"	
	/note="ProfileScan hit to PS0101, P-loop nucleotide binding motif (does not find all)."	
misc_feature	complement(1086..1109)	
	/locus tag="DIP1059"	
misc_feature	/note="ScanRegexp hit to PS00017, ATP/GTP-binding site motif A (P-loop)."	
	complement(1223..2182)	
gene	/locus tag="DIP1060"	
	complement(1223..2182)	
CDS	/locus tag="DIP1060"	
	complement(1223..2182)	
misc_feature	/note="Similar to Streptomyces coelicolor putative iron-siderophore uptake system transmembrane component scs151_26C TR:Q9S214 (EMBL:AL109848) (375 aa) fasta scores: E(): 5e-40, 40.74% id in 324 aa, and to Escherichia coli ferric enterobactin transport system permease protein fepg fepg or B0589 SW:FEPG ECOLI (P23877) (330 aa) fasta scores: E(): 1.8e-45, 44.09% id in 322 aa"	
	/codon start=1	
misc_feature	/transl table=11	
	/product="Putative iron-siderophore uptake system transmembrane component"	
misc_feature	/protein id="CAE49582.1"	
	/db xref="GI:38199915"	
misc_feature	/translations="MRALVLSLLVTAAYSLVPGAGLSTWQLTSTPLAHTVM QRLPRVTGVLGAALAGSLFQSLTRNPLGSPDIIIGFTGAYTVIAFLGWGSG FQATVIGALGGLAVSVVIALSVTRIDGLRIILVGLIGISALSNLNLWIRGELD TALSAASWAGSLNGLRWTIAIPACCVLAIIITPIPLRRLDLVLSIGDDLVAGLRL LGTKLLLVGLVPLVAATVAGPIAFVALASPHARALTSNRPVETSVIGALL IIVADLIGRLFPYQTLPGVLTVTIGGYMLLARGSKENS"	
	complement(2108..2182)	
misc_feature	/locus tag="DIP1060"	
	/note="Signal peptide predicted for DIP1060 by SignalP 2.0. HMM (Signal peptide probability 0.966) with cleavage site probability 0.180 between residues 25 and 26; signal-peptide site"	
misc_feature	complement(1244..2116)	
	/locus tag="DIP1060"	
misc_feature	/note="HMPfam hit to PF01032, FeccD transport family"	
	order(complement(2105..2161), complement(1979..2044), complement(1893..1939), complement(1805..1870), complement(1718..1783), complement(1607..1657), complement(1445..1510), complement(1319..1384), complement(1247..1303))	
misc_feature	/locus tag="DIP1060"	
	/note="9 probable transmembrane helices predicted for DIP1060 by TMHMM2.0"	
misc_feature	complement(1250..1564)	
	/locus tag="DIP1060"	
misc_feature	/note="BlastProDom hit to PD001557, PD001557"	
	complement(2184..3191)	
gene	/locus tag="DIP1061"	
	complement(2184..3191)	
CDS	/locus tag="DIP1061"	
	/note="Similar to Streptomyces coelicolor putative iron-siderophore uptake system transmembrane component scs151_27C TR:Q9S213 (EMBL:AL109848) (348 aa) fasta scores: E(): 8e-36, 36.53% id in 323 aa, and to Escherichia coli ferric enterobactin transport system permease protein fepd or B0590 SW:FEPD ECOLI (P23876) (334 aa) fasta scores: E(): 5.2e-40, 40.37% id in 322 aa"	
	/codon start=1	
	/transl table=11	
misc_feature	/product="Putative iron-siderophore uptake system transmembrane component"	
	/protein id="CAE49584.1"	
misc_feature	/db xref="GI:38199916"	
	/translation="MLQPRVRLPLVITFLVILLCTIIMSLGVGARNITSPDILA ALRHGDPNTHIWRIRPTLTAAGASISLAGVLIQALTRNPLADTVFGINAG AAFITVIGTALGASHISIFALLAGLAGACVVALSNCKSGSDPLRLVLAGVAL NAITIGDGLSLVNPQAFDLKSMVMVNDIAGSVQPNVAGVGLILGLVATCMQI LNALSLGLAITMGASTKRLFTFTAIIVLAASATAAGVITFLGLMVPHARIV GPNNLRVATSLIGIIVISADILGRIIVPGEPFAGVVAFTGAPFLIYAQTAKKE 1"	
misc_feature	complement(3090..3191)	
	/locus tag="DIP1061"	
misc_feature	/note="Signal peptide predicted for DIP1061 by SignalP 2.0. HMM (Signal peptide probability 1.000) with cleavage site probability 0.663 between residues 34 and 35; signal-peptide site"	
	complement(2199..2525)	
misc_feature	/locus tag="DIP1061"	
	/note="BlastProDom hit to PD001557, PD001557"	
misc_feature	complement(2205..3095)	
	/locus tag="DIP1061"	
misc_feature	/note="HMPfam hit to PF01032, FeccD transport family"	
	order(complement(3096..3161), complement(2934..2990), complement(2838..2903), complement(2766..2831), complement(2670..2735), complement(2544..2609), complement(2394..2459), complement(2289..2354), complement(2211..2276))	
misc_feature	/locus tag="DIP1061"	
	/note="9 probable transmembrane helices predicted for DIP1061 by TMHMM2.0"	
gene	3239..4282	
	/locus tag="DIP1062"	
CDS	3239..4282	
	/locus tag="DIP1062"	
misc_feature	/note="Similar to Escherichia coli iron(III) dicitrate-binding periplasmic protein precursor FecB or B4290 SW:FECB ECOLI (F15028) (300 aa) fasta scores: E(): 30.5%; Score 724.4; DB 1; Length 349535; Best Local Similarity 58.2%; Pred. No. 1.2e-166; Matches 1551; Conservative 0; Mismatches 931; Indels 38; Gaps 3;	
	Query Match	
Qy	81	ACTAAGCCGCAATCCCTCGATTCCTGTCATCAACGACGGCTCTGTGAGCTAGTAGAG 140
	38257	ACAGTGGATGTAGTCCCTCGAGGCTGCATCCACTACCGCTTGAGTCAACCCCTTAGGG 38198
Qy	141	ATCTAGATTCACGCGCCATCGTTCGCAATACATCGGTGTCTCAATGGGTATCTCATCGA 200
	38197	CTGGAGATTCACCGCCATCGCTGCCAATACAGCGGAACCTTCGTGCACCTCGCTATCTA 38138
Qy	201	GGAGGATCACTCTCTCTGTTTATAGCATGGAGAGCTGGTTTGGGTTTGGGNAAGTCCCC 260
	38137	GCATCACTACTCAACAGAAATCCAAAAGCGGTGAATATCGGGAGCATTCGCC 38078
Qy	261	AACGAGGCTCGGCAATTCCTCAACAAACCTTCGCCGACGCGGCAATGCGATAGCG 320
	38077	AGCCCAATCCCAACGCGCGCTTCATGATGATCCCTCCGATGAGGAGTCTCATTAACAC 38018
Qy	321	GCCTGCGCCCAACAGGACCATCGAGCGCGCCGCTCCATGATGATGATGATGATGATGAT 380
	38017	GTCGCGCTGCGGACCTTCCAGAGACGCGCTTAATATATATATATATATATATATATAT 37958
Qy	381	TGGGACCGAGGCTAGAGCGGCGATCGGACCCCAATCTAGTTTCCCATCAACCATGTAGG 440
	37957	GAGGCGCAATTCGCAATCTAGGCGATTCGCCCTCCCAATCAACAGGACCTCTATGGTGTAGC 37898
Qy	441	CATCCGCAATGAGGCGGTTTCAATGCGCAAGTGGCGCATGTTCCAAAGTTCTTACTACT 500
	37897	GATCCCAAGAGTTCGGACTTGCCTGCAAAATATCGAATGCACCTAACGCTATGAGT 37838
Qy	501	CACATCCCGCCACGGGATAGTTTCAACGGGTTACCGCTCTCTAAACATCTCCACGCCGA 560

Db 37837 CGCAACAGATACAGGTGCTTTTCCCGCTTACAGCGCAAGGACATCTCCGCGACGCA 37778
 Qy 561 GCAGAGTAATCTGTGGCTTATCTTCCAGCGAGCGTGAAGGTGCTCCACCCCAAG 620
 Db 37777 ACAAGCAACAGAGTGGATCTGCTCTGATACGATATGCAATGTGGGTATTCATCT 37718
 Qy 621 AAGTACCTCGTTGAACACGGGAGGAAACCATGTGATAGCGAATCTGGTGTGATGGGA 680
 Db 37717 CCGCCACATCTCCATCACCGCGGAAACCATGTGATAGCGAATCTGGTGTGATGGGA 37658
 Qy 681 TGGTTAAGCGAGTTCAGCAAGCGCTCCAGATAGTGGCTTATGTTCTGCTGACGA 740
 Db 37657 CAGACATGGAATCGCTCAATCTCTCCCGCACTGAGCGTGTGTTTGGCGTGCAGAA 37598
 Qy 741 ACACATTTTCGGCTGCTGTGACAGGACTTCACCGCTCGGTGTGCTTGGCGGTT 800
 Db 37597 GTGCCATAGCTGCTGCGGATTCACCGATCTCCGCGCATCTGCGGTGATGATGGTG 37538
 Qy 801 GGTGCGGATACCAACACTCGACCGACGATGATGCTCGAGAGCTTTAAAGCGCTGACTCA 860
 Db 37537 ACGTACGGGTACACTACTCGACAGTCTCCCGCTCCAGCGCTTAATCTCTGACTTA 37478
 Qy 861 CGCGGAGGAGGAATGGAAGGGCTAAGAGGGCTTCGAGCTGCTTCATCATGTA 920
 Db 37477 CTGCTGAAGGAGAAATTCCTAATGCAAAAGCGGATCTTCGAGCTGCTCTGATTA 37418
 Qy 921 TTGAGCAAGGTGCTCCAGTGAATGGGTTTCATGAGCTATATTAACCATGTTAAGAA 980
 Db 37417 TTGCGCAAGGTTCTTAATGAAGCGGATTCATGAAGCAATTCATACCAACCTGAAAG 37358
 Qy 981 CCAATCATTTACTTACTTACTTCCATAGTTCAGATGCTGATCATGAAATCTTCATTA 1040
 Db 37357 TCTGTATTTACTTACTTACTTTCAGAGGATATGCTTGTATGATGATGATGCTG 37298
 Qy 1041 CAGTCTGCTTTTGGGGCCAGTCTTTTACTGCTCATCGACCGGAGATGACTGFGA 1100
 Db 37297 CTGGAATTCCTAGTGGTCTGTCTGATGCTGGGCACTGGGCAAAATGGCTTGATTA 37238
 Qy 1101 TTAAACAAGGATTAAGCGGAGGACTCATGCTGCTCTGCTGTTGTTAAATTTCTG 1160
 Db 37237 TTGCTGAAGGATTAAGCGGAGGACTCATGCTGCTCTGCTGTTGTTAAATTTCTG 37178
 Qy 1161 ACGTCTTTTGTGATCGCGGACCTTGGGGGTGATCTTTGTTGATGATGATGATGATG 1220
 Db 37177 ACGTGAATCTCTATTTTGGAGGACCTGCGGTGTGCGGCAATTAAGTCAATGATG 37118
 Qy 1221 TCGTCTCGATATATGCGTGGGGTGGCATGCTTACCTGTTATGTTGTTGCGGTGATGG 1280
 Db 37117 TCGCACTAGTAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37058
 Qy 1281 CAGCGAAGGACCATGACAAACAGGTGGAAGCGCCAGATCATTTGAAGAAACAGAAC 1340
 Db 37057 GTTTAAAGAAAGCATTTAAGCTGCGGCAAGCGCTGCGAGTC---GAGCAAGGAAAC 37001
 Qy 1341 CAACGCTGCCGATGACAGCTTTTGGCGGTTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1400
 Db 37000 CGGTGGCTACGAAACAGTGGCTGACGCTATCATCCGAGTGAATCACTAAACACGCAAA 36941
 Qy 1401 GGTGCGGCTGAGGTGAGCTCGATACGACGCGGTTGGGTAAAGCCATGTTGATGG 1460
 Db 36940 AAGCTAGCCCAAG-----TCAGCGGACGCGCTGCGGTAAACCGAGTCTGCGAG 36890
 Qy 1461 CAATCGTGTGACTGTTGTAACCCGATTTGGAACGCTTTGGAACGCTTTGTTATGCGCG 1520
 Db 36890 CGTTGGCTTTTCACTGGCTTAATCCAGCTGCTTATAGACGCTTTTGTCTATGTTGGAG 36830
 Qy 1521 GGTGCGGCGGAAATACGCGACAGCGAGGTGATTTTCGCGCTGCGGCTGCGGCTGCGG 1580
 Db 36829 GAATCGCAACCAACAGCGTCCGAGCGGCGCTGCGTCTTCGCGCTGCGGCTGCGG 36770
 Qy 1581 CAAGCTGATCTGTTTCCGCTGCTGGTTCGCGGAGGAGCTGTCAGCGCGCTGCT 1640
 Db 36769 CGAGTCTCAGTGGTTCGCTTATTTGATATACCTCCACGCTTTTCTACCGCTGCTAT 36710

Qy 1641 CCAGCCCCAAGGTGTGGCTGTGGATCAACGCTGCTGTGGCAGTTGTGATCAACGCTTGG 1700
 Db 36709 CGCGGCCAGCGTGTGGGATATATCAACATAGCAATCGAATCATCATGATGATCAATGT 36650
 Qy 1701 CCAACAACTGATGTGATGG-----GTAGTTTTCGG 1734
 Db 36649 GCGACGCTCTATCATGCACTAGGCTCAATGCTCTTAAAGGTACCGCGCTTTAATCTGT 36590
 Qy 1735 GGTTTGGAATCGGTGGCTTCCGCCAAATGTTGATGGCGGCTGCTGGGAAATCTCATC 1794
 Db 36589 AACCTTCGAGGCGTAGCACCGCCCAATTAATTAATTCAGCGCTCTTGGCAACAGTGTG 36530
 Qy 1795 GATCGCTCCAACTCGCGCTGAGAAACTCAAGTGTGTGATGAATCAAGCGCTGTGTC 1854
 Db 36529 AATATAATTTGGCTTCTCAACGCTGACTCAAGATTAATTAAGCTCTCCAGATTTGATC 36470
 Qy 1855 CAGTCTCTCAATGACGAGCAACCAATCAATGCACTGCTGCAAGTCACTGCGCGCTTCTG 1914
 Db 36469 CAATCTGCTACCGACGATCTCCGATCAACGCACTGGTTACCGTTGTGGCCCATATC 36410
 Qy 1915 TCCTTGTCTCGCGACGACCCATGCAAGCGCATCTGCGCAAGTCACTGCGCGCTTCTG 1974
 Db 36409 GCCTTGTTCGCGACGACCAACGCAATAGCGCTGTGCGCAAGTGTGGCCAGCTGTAA 36350
 Qy 1975 GCGATGTCAATGAGCTTGGCGACCATATCAATTAATTTGCTACGTTCAATGCTCCCTG 2034
 Db 36349 CGCTAAATCAATTCAGTGCACGACCATGTCAGATTTTTCAGCTTTGAGCATCTCTTGGC 36290
 Qy 2035 CAGGACTTACCTGCTGGTGGCGGGGAACTCTTGGAAATTCATCGAGATATTTGCTCGT 2094
 Db 36289 CAAGGACTTCCCTGCTGCTCGCGTGAATCCACTGCGACGCGCTCGAGATACCTGCTCAGT 36230
 Qy 2095 GAGCGGCTTCCGCAAGTGTGAGAAAGCAATGACGCGACCAATGTTGCGAGCTGA 2154
 Db 36229 TAAAGCGCTTGTGCAAGTGTGCAAAATGCAATAACCCAGTCCGTTATTCGCGCGA 36170
 Qy 2155 GTGCAACAAAGTTCTCAGCTGCTATCGCCCGGTTCTCCACCAAGCAATTAATGAGAAATA 2214
 Db 36169 CTCTAACAGGTTCTCGCATCTTCACTGCTTCTTACCCAGCGGTTCAAGATAGAATA 36110
 Qy 2215 GCTTGTCTGATCAATCAGAGCGGCGAGCGCTCTCCGCGATGAATCAAGCGCTCCG 2274
 Db 36109 GCTCGTGTGATCAATCAGAGCGGCGAGCGCTTCAAGCTAGATTTTCAGATCTTGC 36050
 Qy 2275 TGTGAGCTCTGACGCTGAGAGAAATACCACTGAAAGCGCTTTCAGAGCGAAAT 2334
 Db 36049 GGTTAGTTCAGGCTCCATAGAGAAATTTCCACATAGCGAGCTTTCGCGATGCCAGAT 35990
 Qy 2335 GTCAAGCAATGCTACATGTTCTTCTTCCAAAGGAGTATCT 2374
 Db 35989 GTCTCGCAAGCATATGCAATTTCTTCCAAAGCGTATCT 35950

RESULT 14

AX643028
 LOCUS 711 bp DNA linear PAT 24-FEB-2003
 DEFINITION Sequence 7 from Patent EP126966.

AX643028
 ACCESSION
 AX643028.1 GI:28550158

KEYWORDS
 Corynebacterium glutamicum

ORGANISM
 Corynebacterium glutamicum

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE

1
 Gunji, Y. and Yasueda, H.
 Method for producing L-lysine or L-arginine by using methanol

TITLE
 assimilating bacterium

JOURNAL
 Patent: EP 126966-A 7 18-DEC-2002;

FEATURES
 Ajinomoto Co., Inc. (JP)

Location/Qualifiers
 1..711

source

us-09-105-117k-1.rge

Thu Mar 18 12:30:58 2004

Search completed: March 15, 2004, 22:01:55
Job time : 9515 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 15, 2004, 10:09:44 ; Search time 968.046 Seconds
(without alignments)
10418.122 Million cell updates/sec

Title: US-09-105-117K-1
Perfect score: 2374
Sequence: 1 ccattctgtgaaggtgttac.....ttcttcacaagaggtatct 2374

Scoring table: IDENTITY NUC
Gapop 10_0, Gapext 1.0

Searched: 337863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2374	100.0	2374	2	AAT96816 DNA encod
2	2374	100.0	2374	9	ADB66196 DNA fragm
3	2374	100.0	349980	5	Aah68528 C glutami
C 4	993	41.8	993	4	Aaf71777 Coryneb
C 5	993	41.8	993	4	Aas96096 C. glutam
C 6	870	36.6	870	5	Aah68421 C. glutam
7	822	34.6	822	4	Aaf71779 Coryneb
8	822	34.6	822	4	Aas96098 C. glutam
9	711	29.9	711	8	ACC80941 LysE prot
10	708	29.8	708	5	Aah68420 C glutami
11	698.6	29.4	1568	4	Aah45375 C. thermo
12	696.8	29.4	627	5	Aah68419 C glutami
C 13	627	26.4	627	5	Aaf68077 Coryneb
C 14	485.2	20.4	993	4	ACA29651 Prokaryot
C 15	305.8	12.9	1095	7	ACA29651 Prokaryot
C 16	302.8	12.8	879	7	ACA25567 Prokaryot
C 17	132.6	5.6	1041	7	ACA26879 Prokaryot
C 18	123.8	5.2	897	7	ACA38559 Prokaryot
C 19	122.8	5.2	909	7	ACA38559 Prokaryot
C 20	122.8	5.1	15239	2	AAT33536 BCG delet
C 21	121.2	5.1	912	7	ACA40574 Prokaryot
22	121.2	5.1	110000	4	Continuation (23 o
23	121.2	5.1	110000	4	Continuation (23 o

C 24	116.4	4.9	1041	7	ACA24017 Prokaryot
C 25	114.8	4.8	1107	7	ACA26520 Prokaryot
C 26	108.6	4.6	1038	7	ACA32095 Prokaryot
C 27	108	4.5	1207	9	ADD13356 C. glutam
C 28	107	4.5	1041	7	ACA19150 Prokaryot
C 29	105.2	4.4	999	7	ACA51519 Prokaryot
C 30	105.2	4.4	999	7	ACA50894 Prokaryot
C 31	101	4.3	1200	4	AAF71729 Coryneb
C 32	100.8	4.2	5541	2	AAD55755 Escherich
C 33	97.4	4.1	1035	7	ACA45509 Prokaryot
C 34	93.2	3.9	988	6	ABK72786 Bacillus
C 35	90.4	3.8	349980	6	ABQ81844 Bifidobac
C 36	82.6	3.5	894	4	AAS2558 E. coli D
C 37	82.6	3.5	894	7	ACA32622 Prokaryot
C 38	81.4	3.4	999	7	ACA24884 Prokaryot
C 39	79.8	3.4	909	7	ACA53702 Prokaryot
C 40	79.4	3.3	23128	4	AAS9552 Propionib
C 41	79.4	3.3	23128	7	ACF64481 Propionib
C 42	78.2	3.3	990	7	ACA35500 Prokaryot
C 43	78	3.3	1077	5	Aah66360 C glutami
C 44	77.6	3.3	903	7	ACF71729 Photornab
C 45	77.6	3.3	110000	7	Continuation (50 o

ALIGNMENTS

RESULT 1
ID AAT96816 standard; DNA; 2374 BP.
XX
AC AAT96816;
DT 12-MAR-1998 (first entry)
XX
DE DNA encoding LysG, LysE and ORF3 from Corynebacterium glutamicum.
KW LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;
KW Microbial production; amino acid; animal feed additive; ds.
XX
OS Corynebacterium glutamicum.

XX Key Location/Qualifiers
FH complement (82. .954)
FT /*tag= a
FT /label= LysG
FT 1016. .1726
FT /*tag= b
FT /label= LysE
FT complement (1723. .2373)
FT /*tag= c
FT /label= orf3

DE19548222-N1.

26-JUN-1997.

22-DEC-1995; 95DE-01048222.

22-DEC-1995; 95DE-01048222.

(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

Vrljic M, Eggeling L, Sahn H;

WPI, 1997-333867/31.

P-PSDB; AAN37714, AAN37715, AAN37716.

Increasing microbial production of amino acids, especially lysine - by
improving export carrier activity or corresponding gene expression, also
new export and regulatory genes from Corynebacterium.

Claim 23 and 26; Page: 16pp; German.

QY 781 TTGCGTGTCTTGGCGCGTGGGTGGCGATACCAACACTCGACCCACGATGATCTCGAG 840
Db 781 TTGCGTGTCTTGGCGCGTGGGTGGCGATACCAACACTCGACCCACGATGATCTCGAG 840
QY 841 AGCTTTAAACCGCTGACTACCCGCGAGGGGAAATGGAAGGGCTAAGAGGCGCCCTTC 900
Db 841 AGCTTTAAACCGCTGACTACCCGCGAGGGGAAATGGAAGGGCTAAGAGGCGCCCTTC 900
QY 901 GAAGTGCCTTCATCAATGATTTGAGAGCAAGTGTCCAGTTGAATGGGGTTCATGAAGCT 960
Db 901 GAAGTGCCTTCATCAATGATTTGAGAGCAAGTGTCCAGTTGAATGGGGTTCATGAAGCT 960
QY 961 ATATTAAACCATGTTAAGAACCAATCAATTTACTTTAAGTACTTCCATAGGTTCAGATGCT 1020
Db 961 ATATTAAACCATGTTAAGAACCAATCAATTTACTTTAAGTACTTCCATAGGTTCAGATGCT 1020
QY 1021 GATCATGGAATCTTCATTAAGTCTGCTTTGGGGGCGAGTCTTTTACTGTCCATCGG 1080
Db 1021 GATCATGGAATCTTCATTAAGTCTGCTTTGGGGGCGAGTCTTTTACTGTCCATCGG 1080
QY 1081 ACCGCAGATGTACTGSGTGAATTAACAAGGAATTAAGCGGAGGACTCAATCGGTTCT 1140
Db 1081 ACCGCAGATGTACTGSGTGAATTAACAAGGAATTAAGCGGAGGACTCAATCGGTTCT 1140
QY 1141 TCTGCTGTGTTAAATCTGAGCTCTTTTGTTCATCGCGGACCTTGGCGGTTGATCT 1200
Db 1141 TCTGCTGTGTTAAATCTGAGCTCTTTTGTTCATCGCGGACCTTGGCGGTTGATCT 1200
QY 1201 TTTGTCCAATGCCGCGCGATCGTGTTCGATATTATCGCTGGGGTGGCATCTGCTTACCT 1260
Db 1201 TTTGTCCAATGCCGCGCGATCGTGTTCGATATTATCGCTGGGGTGGCATCTGCTTACCT 1260
QY 1261 GTTATGTTTGGCGTCTGCGAGCAAGAGCGCCATGACACACAGAGTGGAGCGCCACA 1320
Db 1261 GTTATGTTTGGCGTCTGCGAGCAAGAGCGCCATGACACACAGAGTGGAGCGCCACA 1320
QY 1321 GATCATTTGAAGAACCAACACCGTGGCGGATGACACGCTTTGGCGGTTGGCGGCT 1380
Db 1321 GATCATTTGAAGAACCAACACCGTGGCGGATGACACGCTTTGGCGGTTGGCGGCT 1380
QY 1381 GGCACACTGACACCGGCGGCGGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1440
Db 1381 GGCACACTGACACCGGCGGCGGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1440
QY 1441 GGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGTTGAACCCGAATGGTATTGGA 1500
Db 1441 GGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGTTGAACCCGAATGGTATTGGA 1500
QY 1501 CGCGTTGTGTTTATCGCGCGCTGCGCGGCAATGACGCGACACCGGAGCGTGGATTTT 1560
Db 1501 CGCGTTGTGTTTATCGCGCGCTGCGCGGCAATGACGCGACACCGGAGCGTGGATTTT 1560
QY 1561 CGCGCTGGCGGCTTGGCGGAGCTGATCTGGTTCCGCTGGGTGGGTTCGGCGGAGC 1620
Db 1561 CGCGCTGGCGGCTTGGCGGAGCTGATCTGGTTCCGCTGGGTGGGTTCGGCGGAGC 1620
QY 1621 AGCATTTGACGCGCTGTCTGAGCGCCCAAGGTGGCGTGGATCAACGCTCGTGGCG 1680
Db 1621 AGCATTTGACGCGCTGTCTGAGCGCCCAAGGTGGCGTGGATCAACGCTCGTGGCG 1680
QY 1681 AGTTGTGATGACCGGATTTGGCCATCAACTGATGTTGATGGGTAGTTTTCGGCGGTTT 1740
Db 1681 AGTTGTGATGACCGGATTTGGCCATCAACTGATGTTGATGGGTAGTTTTCGGCGGTTT 1740
QY 1741 GGAATCGGTGGCTTCGCGCAAAATGTTGATGCGCGGCTGCTGGGAAATCTCATCGATCGC 1800
Db 1741 GGAATCGGTGGCTTCGCGCAAAATGTTGATGCGCGGCTGCTGGGAAATCTCATCGATCGC 1800
QY 1801 CTCCAACTCGCGCTGAGAAACTCAAGTTGTTGAGTGAATCAAGCTGTTGTCAGCTG 1860
Db 1801 CTCCAACTCGCGCTGAGAAACTCAAGTTGTTGAGTGAATCAAGCTGTTGTCAGCTG 1860
QY 1861 CTCCAACTGAGAAACCAATCAATGCACTGCTACGCTATCGCGCGGCTACTCTCCTTG 1920

Db 1861 CTCCAACTGAGAAACCAATCAATGCACTGCTACGCTATCGCGCGGCTACTCTCCTTG 1920
QY 1921 CTCGCGAGCACCATGCAAGCGCCATCTGCGAAGTGACTGCGCGGTTCTTGGGCGAT 1980
Db 1921 CTCGCGAGCACCATGCAAGCGCCATCTGCGAAGTGACTGCGCGGTTCTTGGGCGAT 1980
QY 1981 GTCATTTAGCTTGGCGGACCATATCAATATTTGTTCACGTTCAACATGCGCTCAGACGGA 2040
Db 1981 GTCATTTAGCTTGGCGGACCATATCAATATTTGTTCACGTTCAACATGCGCTCAGACGGA 2040
QY 2041 CTTACCTCTGCTGGCGGGGAAACCTCTGGAATTTCCATGAGATATTTGTCGCTGAGCAG 2100
Db 2041 CTTACCTCTGCTGGCGGGGAAACCTCTGGAATTTCCATGAGATATTTGTCGCTGAGCAG 2100
QY 2101 GCGCTGGCGAAGTGTGAGAAAGCAATGACGCAAGACCATTTGTTGGAGCTGACTGCAA 2160
Db 2101 GCGCTGGCGAAGTGTGAGAAAGCAATGACGCAAGACCATTTGTTGGAGCTGACTGCAA 2160
QY 2161 CAAGTTCTCAGCTCATCGCCCGGTTCTCTCCACCAACGATTAATGATGGAATAGCTTG 2220
Db 2161 CAAGTTCTCAGCTCATCGCCCGGTTCTCTCCACCAACGATTAATGATGGAATAGCTTG 2220
QY 2221 CTGATGATCAGAGCGGCGAGCCCTCTCGSCCATGAATCAGCGCGCTCCGCTGAG 2280
Db 2221 CTGATGATCAGAGCGGCGAGCCCTCTCGSCCATGAATCAGCGCGCTCCGCTGAG 2280
QY 2281 CTCTGGAACCTAGGAAGAAATACCCACGTAAGAGCCTTTCCAGACGCAAAATGTCAG 2340
Db 2281 CTCTGGAACCTAGGAAGAAATACCCACGTAAGAGCCTTTCCAGACGCAAAATGTCAG 2340
QY 2341 CAATGCTGATGCTTTCTTCCAAAGGATATCT 2374
Db 2341 CAATGCTGATGCTTTCTTCCAAAGGATATCT 2374

RESULT 3

AAH68528/c

ID AAH68528 standard; DNA; 349980 BP.

XX AC AAH68528;

XX DT 26-SEP-2001 (first entry)

XX DE C glutamicum coding sequence fragment SEQ ID NO: 7063.

XX KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX OS organic acid synthesis; ds.

XX OS Corynebacterium glutamicum.

XX PN EP1108790-A2.

XX PD 20-JUN-2001.

XX PF 18-DEC-2000; 2000EP-00127688.

XX PR 16-DEC-1999; 99JP-00377484.

XX PR 07-APR-2000; 2000JP-00159162.

XX PR 03-AUG-2000; 2000JP-00280988.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX DR WPI; 2001-376931/40.

XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying

XX PT mutation point of a gene, measuring expression of a gene, analyzing

XX PT expression profile or pattern of a gene and identifying homologous gene.

XX PS Disclosure; SEQ ID NO 7063; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium, and identifying a homologue of a gene derived from
CC corynebacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office

XX SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;
Query Match 100.0%; Score 2374; DB 5; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATTTGCTGAAGGTGTTACTCTGCTGGCCCAATTCCTGGGGCGAAGTGAATAAC 60
DB 129968 CCATTTGCTGAAGGTGTTACTCTGCTGGCCCAATTCCTGGGGCGAAGTGAATAAC 129909

QY 61 CCTGAACCTTTTCAGAGTAACCTAAGCGCGCAATCCCTCGATGCTGCATCAACGACGCG 120
DB 129908 CCTGAACCTTTTCAGAGTAACCTAAGCGCGCAATCCCTCGATGCTGCATCAACGACGCG 129849

QY 121 GTCTGTGAGTCTAGTAGAGATCTAGATTCAGAGCGCATCTGCTGCTCCCAATACATCGGTGT 180
DB 129848 GTCTGTGAGTCTAGTAGAGATCTAGATTCAGAGCGCATCTGCTGCTCCCAATACATCGGTGT 129789

QY 181 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTGTAGCATGGAGGACGCTTG 240
DB 129788 GTCAATGGGTATCTCATCGAGGAGATCACTTCTCTGCTTTTGTAGCATGGAGGACGCTTG 129729

QY 241 GGTTCGGGAAGAATCCCCAACCAAGGCTCGGGAATTCCTCACCAGAACTTCGCG 300
DB 129728 GGTTCGGGAAGAATCCCCAACCAAGGCTCGGGAATTCCTCACCAGAACTTCGCG 129669

QY 301 CGAGCGCAATAGTATAGCGCTCGCCGCCACAGAGCAATCGACGCGCCGCTCCAGTC 360
DB 129668 CGAGCGCAATAGTATAGCGCTCGCCGCCACAGAGCAATCGACGCGCCGCTCCAGTC 129609

QY 361 ACGGTCTTGAAGCACATCTTTGGACCGAAGCGTAAAGCGGATCGACGCCCAATCTAG 420
DB 129608 ACGGTCTTGAAGCACATCTTTGGACCGAAGCGTAAAGCGGATCGACGCCCAATCTAG 129549

QY 421 TTTCCCATCAACCATAGGCAATCCGCAATGAGGGGTTCGATGSCCAAGTGGCGCAT 480
DB 129548 TTTCCCATCAACCATAGGCAATCCGCAATGAGGGGTTCGATGSCCAAGTGGCGCAT 129489

QY 481 GGTTCAGAGTCTTACTACTTACATCCGCGCACAGGATTTAGCTTTACGGGTTACCGCTCC 540
DB 129488 GGTTCAGAGTCTTACTACTTACATCCGCGCACAGGATTTAGCTTTACGGGTTACCGCTCC 129429

QY 541 TAAACATCTCCACGCGCGAGCAAGATAATGTGCGCTTCATCTTCCAGCGAGCGGT 600
DB 129428 TAAACATCTCCACGCGCGAGCAAGATAATGTGCGCTTCATCTTCCAGCGAGCGGT 129369

QY 601 GAGCGTGTCTCCACCCAGAGCTTACCTGTTGACACGAGGAGAAACCATGTGATAG 660
DB 129368 GAGCGTGTCTCCACCCAGAGCTTACCTGTTGACACGAGGAGAAACCATGTGATAG 129309

QY 661 CGAATCTGCTGTGATGCGATGTTAAACGGATTTAGCAAGGCGCTCCAGATAGTGGCG 720
DB 129308 CGAATCTGCTGTGATGCGATGTTAAACGGATTTAGCAAGGCGCTCCAGATAGTGGCG 129249

QY 721 TTTAGTTTCTGCTGACGCAACACATTTTCGCGCTGTGCTGCAAGAGCTTACCGCG 780
DB 129248 TTTAGTTTCTGCTGACGCAACACATTTTCGCGCTGTGCTGCAAGAGCTTACCGCG 129189

QY 781 TTCGGTGTCTTGGCCGTTGGGTGGCGGATACCACTCGACCCACGCTGCTCGAG 840

DB 129188 TTCCGTTGCTTTGGCGCGTGGGTGGCGGATACCAACTCGACCCACGATGATGCTCGAG 129129

QY 841 AGCTTTAAGCGCTGACTCACCGCGAGGGGGAAATGGAAGGGCTAAGGAGCGCCTTC 900

DB 129128 AGCTTTAAGCGCTGACTCACCGCGAGGGGGAAATGGAAGGGCTAAGGAGCGCCTTC 129069

QY 901 GAAGCTGCTTTCATCAATGATTTGAGAGCAAGTGTCCAGTTGATGGGTTCATGAAGCT 960

DB 129068 GAAGCTGCTTTCATCAATGATTTGAGAGCAAGTGTCCAGTTGATGGGTTCATGAAGCT 129009

QY 961 ATATTAACCATGTTTGAACCAATCTTTTACTTAAGTACTTCCATAGTCTACGATGT 1020

DB 129008 ATATTAACCATGTTTGAACCAATCTTTTACTTAAGTACTTCCATAGTCTACGATGT 128949

QY 1021 GATCATGGAATCTTTCAATACAGGTCTGCTTTTGGGGGCGAGTCTTTTACTGTCATCGG 1080

DB 128948 GATCATGGAATCTTTCAATACAGGTCTGCTTTTGGGGGCGAGTCTTTTACTGTCATCGG 128889

QY 1081 ACCGAGAAATGACTGTGTGATTAACAGGAATTAAGCGGAGAGACTCATTTGCGGTCT 1140

DB 128888 ACCGAGAAATGACTGTGTGATTAACAGGAATTAAGCGGAGAGACTCATTTGCGGTCT 128829

QY 1141 TCTGCTGTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACTTTGGCGGTGATCT 1200

DB 128828 TCTGCTGTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCGGCACTTTGGCGGTGATCT 128769

QY 1201 TTTCTCCAAATCGCGCGCGATCGTCTCGATTAATATGCGTGGGTGGCATCGCTTACCT 1260

DB 128768 TTTCTCCAAATCGCGCGCGATCGTCTCGATTAATATGCGTGGGTGGCATCGCTTACCT 128709

QY 1261 GTTATGTTGGCTCATGCGGAGAAAGACGCGCATGACAAACAAGTGGAGGCGCACCA 1320

DB 128708 GTTATGTTGGCTCATGCGGAGAAAGACGCGCATGACAAACAAGTGGAGGCGCACCA 128649

QY 1321 GATCATTTGAAGAAACAGAACCAACCGTCCCATGACACGCTTTGGCGGTTCGGCGGT 1380

DB 128648 GATCATTTGAAGAAACAGAACCAACCGTCCCATGACACGCTTTGGCGGTTCGGCGGT 128589

QY 1381 GGCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGGATTAAGCAGCGGGTTG 1440

DB 128588 GGCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGGATTAAGCAGCGGGTTG 128529

QY 1441 GGTAAAGCCCATGTTGATGCAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500

DB 128528 GGTAAAGCCCATGTTGATGCAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128469

QY 1501 CGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560

DB 128468 CGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128409

QY 1561 CGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620

DB 128408 CGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128349

QY 1621 AGCATTTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680

DB 128348 AGCATTTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128289

QY 1681 AGTTGTGATGACCGCATTTGCGCATCAAACTCATGTTGATGGGTGATGTTTGGCGGT 1740

DB 128288 AGTTGTGATGACCGCATTTGCGCATCAAACTCATGTTGATGGGTGATGTTTGGCGGT 128229

QY 1741 GGAATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800

DB 128228 GGAATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128169

QY 1801 CTCCTACTCGCGCTGAGAACTCCAGTGTGCTGATGATCAAGGCTGTTGCTGCTGCTGCTG 1860

DB 128168 CTCCTACTCGCGCTGAGAACTCCAGTGTGCTGATGATCAAGGCTGTTGCTGCTGCTGCTG 128109

QY 1861 CTCCTACTCGAGAGCAACCAATCAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920

QY	62	CTGAACCTTTTTCAGAGTAACCTAAGGCGCGCAATCCCTCGATTGCTGCTCATCAACGCGCG	121
Db	993	CTGAACCTTTTTCAGAGTAACCTAAGGCGCGCAATCCCTCGATTGCTGCTCATCAACGCGCG	934
QY	122	TCTGTGAGTCTAGCTAGAGTCTAGATCCAGCGCCCATCGTTCGCAATACATCGGTGTG	181
Db	933	TCTGTGAGTCTAGCTAGAGTCTAGATCCAGCGCCCATCGTTCGCAATACATCGGTGTG	874
QY	182	TCAATGGGTATCTATCGAGGAGGATCACTTCTCTGCTTTTAGCATGGGAGCAGCTTGG	241
Db	873	TCAATGGGTATCTATCGAGGAGGATCACTTCTCTGCTTTTAGCATGGGAGCAGCTTGG	814
QY	242	GTTCGGGAAGATCCCAACCAAGGCTCGGCGAATTCCTCAGCAAACTTCGCC	301
Db	813	GTTCGGGAAGATCCCAACCAAGGCTCGGCGAATTCCTCAGCAAACTTCGCC	754
QY	302	GACGGGAACATGATAGCGCTCGGCGCCACAGGACCATCGAGCGCCGCTCCAGGTCA	361
Db	753	GACGGGAACATGATAGCGCTCGGCGCCACAGGACCATCGAGCGCCGCTCCAGGTCA	694
QY	362	CGGTCTTGAACACATCTTTGGGACCGCAAGCGTAAAGCGGCAATCGAGCCCAATCTAGT	421
Db	693	CGGTCTTGAACACATCTTTGGGACCGCAAGCGTAAAGCGGCAATCGAGCCCAATCTAGT	634
QY	422	TTCCCATCAACCATGTAGGCATCCCGCAATGAGGGGTTGCAATGCGCAGTGGCGCATG	481
Db	633	TTCCCATCAACCATGTAGGCATCCCGCAATGAGGGGTTGCAATGCGCAGTGGCGCATG	574
QY	482	GTTCAGATTTTACTTACTTCAATCCCGCCAGCGGATAGCTTCAAGCGGTTACCGCTCT	541
Db	573	GTTCAGATTTTACTTACTTCAATCCCGCCAGCGGATAGCTTCAAGCGGTTACCGCTCT	514
QY	542	AAACATCTCCAGCGCGCAGCAAGATATGTGTCGCTTCACTTCCAGCGCAGGCTG	601
Db	513	AAACATCTCCAGCGCGCAGCAAGATATGTGTCGCTTCACTTCCAGCGCAGGCTG	454
QY	602	AGCTGTGCTCCACCCCAAGAGTACTCTGTTGAACCGGAGGAGAACCATGTGATAGC	661
Db	453	AGCTGTGCTCCACCCCAAGAGTACTCTGTTGAACCGGAGGAGAACCATGTGATAGC	394
QY	662	GAATCTGCTGTATGCGCATGTTAAACGGGATTTCAAGCGGCTCCAGATAGTTCGCT	721
Db	393	GAATCTGCTGTATGCGCATGTTAAACGGGATTTCAAGCGGCTCCAGATAGTTCGCT	334
QY	722	TTAGTTTCTGCTTGCAGCAACCACTTTCCGCGCTGTTGCAACAGGACTTCACCGCT	781
Db	333	TTAGTTTCTGCTTGCAGCAACCACTTTCCGCGCTGTTGCAACAGGACTTCACCGCT	274
QY	782	TCGGTTGCTTGGCGGTTGGGTGGCGATACCAACCTCCAGCCAGTGTGCTCGAGA	841
Db	273	TCGGTTGCTTGGCGGTTGGGTGGCGATACCAACCTCCAGCCAGTGTGCTCGAGA	214
QY	842	GCTTTAAGCGCTGCTACACCGCGAGGGGAAATCGAAAGGCTTAAGGAGCGCTTCG	901
Db	213	GCTTTAAGCGCTGCTACACCGCGAGGGGAAATCGAAAGGCTTAAGGAGCGCTTCG	154
QY	902	AACTGCTCTTCAATCATGATTTGAGAGCAAGTGTCAAGTTGAATGGGGTTTCAGACTA	961
Db	153	AACTGCTCTTCAATCATGATTTGAGAGCAAGTGTCAAGTTGAATGGGGTTTCAGACTA	94
QY	962	TATTAACCATGTTAAGACCAATCATTTTACTTATAGTACTTCCATAGTTCAGATGGT	1021
Db	93	TATTAACCATGTTAAGACCAATCATTTTACTTATAGTACTTCCATAGTTCAGATGGT	34
QY	1022	ATCATGGAAATCTTCAATCAGCTCTGCTTTTG	1054
Db	33	ATCATGGAAATCTTCAATCAGCTCTGCTTTTG	1

RESULT 5
AAS96096/c
ID AAS96096 standard; DNA; 993 BP.
XX

Thu Mar 18 12:30:59 2004

693 CGGTCTTGAAGCACAATCTTTGGACCGAAGCGTAAGACGGGCATCGACGCCCAATCTAGT 634
422 TTCCCATCAACCATGTAGGCATCCCGCAATGAGGGGGTTGCAATGGCCCAAGTGGCGCATG 481
633 TTCCCATCAACCATGTAGGCATCCCGCAATGAGGGGGTTGCAATGGCCCAAGTGGCGCATG 574
482 GTTCCAGTTCTACTTACATCCCGCCAGCGGATAGCTTACCGGTTTACCGTCTCT 541
573 GTTCCAGTTCTACTTACATCCCGCCAGCGGATAGCTTACCGGTTTACCGTCTCT 514
542 AAAACATCTCCACCGCCGACGAGGATATGTGTGGCTTCATCTTCCAGCGGAGGTG 601
513 AAAACATCTCCACCGCCGACGAGGATATGTGTGGCTTCATCTTCCAGCGGAGGTG 454
602 AGCGTTGCTCCACCCCAAGCATCTCTGTTGAACAGCGGAGGAAACCATGTGATAGC 661
453 AGCGTTGCTCCACCCCAAGCATCTCTGTTGAACAGCGGAGGAAACCATGTGATAGC 394
662 GAATCTGCTGTGATGGCATGTTAAACGGGATTTGACGAGGGGTCCAGATAGTTGGCT 721
393 GAATCTGCTGTGATGGCATGTTAAACGGGATTTGACGAGGGGTCCAGATAGTTGGCT 334
722 TTAGTTTCTGTTGACGACCAACCATTTCCGCGTGTTCACAGGACTTCCACCGCT 781
333 TTAGTTTCTGTTGACGACCAACCATTTCCGCGTGTTCACAGGACTTCCACCGCT 274
782 TCGGTTGCTTTGGCCGTTGGGTGGCGGATACCAACACTCGACCCAGCTGATGTCGAGA 841
273 TCGGTTGCTTTGGCCGTTGGGTGGCGGATACCAACACTCGACCCAGCTGATGTCGAGA 214
842 GCTTTAAACGCGTCACTCACCGCCGAGGGGGAATGGAAGGGCTTAAGAGGGCGCTTCG 901
213 GCTTTAAACGCGTCACTCACCGCCGAGGGGGAATGGAAGGGCTTAAGAGGGCGCTTCG 154
902 AAGCTGCTTTCATCAATGATGAGAGCAAGTGTCCAGTTGAATGGGTTTCATGAAGCTA 961
153 AAGCTGCTTTCATCAATGATGAGAGCAAGTGTCCAGTTGAATGGGTTTCATGAAGCTA 94
962 TATTAAACCATGTTAAGAACCAATCATTTTAACTTAACTTCCATAGTTCAGATGGTG 1021
93 TATTAAACCATGTTAAGAACCAATCATTTTAACTTAACTTCCATAGTTCAGATGGTG 34
1022 ATCATGGAATCTTCAATACAGGTCGCTTTTG 1054
33 ATCATGGAATCTTCAATACAGGTCGCTTTTG 1

RESULT 6
AAH68421/c
ID AAH68421 standard; DNA; 870 BP.
XX AC AAH68421;
XX AC
XX AC
DT 26-SEP-2001 (first entry)
XX DE C glutamicum coding sequence fragment SEQ ID NO: 3456.
XX DE C glutamicum coding sequence fragment SEQ ID NO: 3456.
XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX KW organic acid synthesis; ds.
XX OS Corynebacterium glutamicum.
XX PN EP1108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-00127688.
XX PR 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-00159162.
XX PR 03-AUG-2000; 2000JP-00280968.
XX

(KYOW) KYOWA HAKKO KOGYO KK.
Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
Tateishi N, Senoh A, Ikeda M, Ozaki A;
WPI; 2001-376931/40.
P-PSDB; AAG93202.
Novel polynucleotides derived from Coryneform bacteria, for identifying
mutation point of a gene, measuring expression of a gene, analyzing
expression profile or pattern of a gene and identifying homologous gene.
Claim 1; SEQ ID NO 3456; 246pp + Sequence Listing; English.
The present invention provides a number of nucleotide and protein
sequences from the Coryneform bacterium Corynebacterium glutamicum. These
sequences are useful for identifying the mutation point of a gene derived from a
mutant of coryneform bacterium, measuring expression amount and analysing
the expression profile or expression pattern of a gene derived from
Coryneform bacterium, and identifying a homologue of a gene derived from
Coryneform bacterium. Coryneform bacteria are useful for producing amino
acids; nucleic acids, vitamins, saccharides and organic acids,
particularly L-lysine. The present sequence is a nucleic acid described
in the exemplification of the invention. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from the European Patent Office
Sequence 870 BP; 192 A; 236 C; 245 G; 197 T; 0 U; 0 Other;
Query Match 36.6%; Score 870; DB 5; Length 870;
Best Local Similarity 100.0%; Pred. No. 1.3e-260;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 85 AGGCCCAATCCCTCGATCTGCTGCATCAACGACGGCTCTGTGAGTCTAGTAGATCT 144
DB 870 AGGCCCAATCCCTCGATCTGCTGCATCAACGACGGCTCTGTGAGTCTAGTAGATCT 811
QY 145 AGATTCAGGGCCATCGTTGCCAATACATCGTGTCAATGGGTATCTCATCGAGGAG 204
DB 810 AGATTCAGGGCCATCGTTGCCAATACATCGTGTCAATGGGTATCTCATCGAGGAG 751
QY 205 GATCAGTCTCTGCTTTTAGCATGGGAGCAGTCTGGGTTTCGGGAGAGTCCCAACC 264
DB 750 GATCAGTCTCTGCTTTTAGCATGGGAGCAGTCTGGGTTTCGGGAGAGTCCCAACC 691
QY 265 AAGGCTCGGGCAATGGCTCCAAAACCTTCGCCGACGGGCAATGATAGCGGCT 324
DB 690 AAGGCTCGGGCAATGGCTCCAAAACCTTCGCCGACGGGCAATGATAGCGGCT 631
QY 325 GGGCCCAACAGGACCATCGACGGCCCGTCCAGGTCAAGGTCTTTGAAGCACAATCTTTGG 384
DB 630 GGGCCCAACAGGACCATCGACGGCCCGTCCAGGTCAAGGTCTTTGAAGCACAATCTTTGG 571
QY 385 ACCGAAGCGTAAGACGGGCATCGCAGCCCAATCTAGTTTCCATCAACCATGAGGATC 444
DB 570 ACCGAAGCGTAAGACGGGCATCGCAGCCCAATCTAGTTTCCATCAACCATGAGGATC 511
QY 445 CCGCAATGAGGGGTTGCAATGGCCCAAGTGGCGCATGTTCCAGTTCTACTACTTACA 504
DB 510 CCGCAATGAGGGGTTGCAATGGCCCAAGTGGCGCATGTTCCAGTTCTACTACTTACA 451
QY 505 TCCCGCCAGGGATTAGTTCAGCGGTTTACCGGTCTCTAAACATCTCCAGCGCGAGCA 564
DB 450 TCCCGCCAGGGATTAGTTCAGCGGTTTACCGGTCTCTAAACATCTCCAGCGCGAGCA 391
QY 565 GGATAATGTGGCTTTCATCTTCCAGCGGAGGATGAGGTTGCTCCACCCCAAGAGC 624
DB 390 GGATAATGTGGCTTTCATCTTCCAGCGGAGGATGAGGTTGCTCCACCCCAAGAGC 331
QY 625 TACCTGTTGAACACGGGAGGAAACCATGTGATAGGATCTGCGTTGATGGCGATGCT 684
DB 330 TACCTGTTGAACACGGGAGGAAACCATGTGATAGGATCTGCGTTGATGGCGATGCT 271

QY 685 TAACGGGATTCAGCAAGCGCTCCAGATAGTTGGCTTTAGTTCTTCTGCTTCAGCAACAC 744
Db 270 TAACGGGATTCAGCAAGCGCTCCAGATAGTTGGCTTTAGTTCTTCTGCTTCAGCAACAC 211
QY 745 CATTTTCCGGCTGCTTCAGCAAGCGCTTCAGATAGTTGGCTTTAGTTCTTCTGCTTCAGCAACAC 804
Db 210 CATTTTCCGGCTGCTTCAGCAAGCGCTTCAGATAGTTGGCTTTAGTTCTTCTGCTTCAGCAACAC 151
QY 805 GCGGATACCAACACTTCAGCAAGCGCTTCAGATAGTTGGCTTTAGTTCTTCTGCTTCAGCAACAC 864
Db 150 GCGGATACCAACACTTCAGCAAGCGCTTCAGATAGTTGGCTTTAGTTCTTCTGCTTCAGCAACAC 91
QY 865 CGAGGGGAAATCGAAGCGCTTCAGCAAGCGCTTCAGATAGTTGGCTTTAGTTCTTCTGCTTCAGCAACAC 924
Db 90 CGAGGGGAAATCGAAGCGCTTCAGCAAGCGCTTCAGATAGTTGGCTTTAGTTCTTCTGCTTCAGCAACAC 31
QY 925 GAGCAAGTGTCCAGTTGAATGGGTTTCAT 954
Db 30 GAGCAAGTGTCCAGTTGAATGGGTTTCAT 1

RESULT 7
ID AAF71779 standard; DNA; 822 BP.
AC AAF71779;
DT 30-APR-2001 (first entry)
XX Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:53.
DE
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB000923.
XX
PR 25-JUN-1999; 99US-0141031P.
PR 01-JUL-1999; 99DE-01030476.
PR 02-JUL-1999; 99US-0142101P.
PR 08-JUL-1999; 99DE-01031415.
PR 08-JUL-1999; 99DE-01031418.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031435.
PR 08-JUL-1999; 99DE-01031443.
PR 08-JUL-1999; 99DE-01031453.
PR 08-JUL-1999; 99DE-01031457.
PR 08-JUL-1999; 99DE-01031465.
PR 08-JUL-1999; 99DE-01031478.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031541.
PR 08-JUL-1999; 99DE-01031573.
PR 08-JUL-1999; 99DE-01031592.
PR 08-JUL-1999; 99DE-01031632.
PR 08-JUL-1999; 99DE-01031634.
PR 08-JUL-1999; 99DE-01031636.
PR 08-JUL-1999; 99DE-01032125.
PR 08-JUL-1999; 99DE-01032126.
PR 08-JUL-1999; 99DE-01032130.
PR 08-JUL-1999; 99DE-01032186.
PR 08-JUL-1999; 99DE-01032206.

PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032228.
PR 09-JUL-1999; 99DE-01032229.
PR 09-JUL-1999; 99DE-01032230.
PR 14-JUL-1999; 99DE-01032232.
PR 14-JUL-1999; 99DE-01032236.
PR 14-JUL-1999; 99DE-01032238.
PR 14-JUL-1999; 99DE-01033004.
PR 14-JUL-1999; 99DE-01033005.
PR 14-JUL-1999; 99DE-01033006.
PR 12-AUG-1999; 99US-0148613P.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99DE-01040765.
PR 27-AUG-1999; 99DE-01040766.
PR 27-AUG-1999; 99DE-01040832.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041379.
PR 31-AUG-1999; 99DE-01041380.
PR 31-AUG-1999; 99DE-01041394.
PR 31-AUG-1999; 99DE-01041396.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042077.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042124.
PR 03-SEP-1999; 99DE-01042129.
PR 09-MAR-2000; 2000US-0187970P.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
XX WPI; 2001-137957/14.
XX P-PSDB; AAB79660.
XX
PT Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway
PT proteins, useful for producing fine chemicals in microorganisms,
PT including organic acids, nonproteinogenic amino acids, and purine and
PT pyrimidine bases.
XX
PS Claim 3; Page 233-234; 1737pp; English.
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
CC nucleic acids are useful for the production of fine chemicals in
CC microorganisms, including organic acids, nonproteinogenic amino acids,
CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
CC vitamins, cofactors, polyketides and enzymes
XX
SQ Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;
Query Match 34.6%; Score 822; DB 4; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.2e-245;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 925 GAGCAAGTGTCCAGTTGAATGGGTTTCATGAGTGTATATTAACCATGTTAAGAACCA 984
Db 1 GAGCAAGTGTCCAGTTGAATGGGTTTCATGAGTGTATATTAACCATGTTAAGAACCA 60
QY 985 TCATTTTACTTAAGTACTTCCATAGTTCAGATGGTGTATGCAATCTTCATTACAGG 1044
Db 61 TCATTTTACTTAAGTACTTCCATAGTTCAGATGGTGTATGCAATCTTCATTACAGG 120
QY 1045 TCTGCTTTTGGGGCCAGTCTTTTACTGTCATCGCACCGCAGAGTGTACTGTTGATTAA 1104
Db 121 TCTGCTTTTGGGGCCAGTCTTTTACTGTCATCGCACCGCAGAGTGTACTGTTGATTAA 180
QY 1105 ACAAGGATTAAGCGCGAAGGACTCAITGGGTTCTTCGTGTTGTTTAACTTCAGCT 1164

Db 181 ACAGGAATTAAAGCGGAGGACTCATTGGGTTCTTCTCGTGTTAAATTTCTGAGCT 240
QY 1165 CTTTGTGTTCAATCGCGGCACTTGGGGTTGATCTTTTGTCCAAATCGCGCGGATCGT 1224
Db 241 CTTTGTGTTCAATCGCGGCACTTGGGGTTGATCTTTTGTCCAAATCGCGCGGATCGT 300
QY 1225 GCTCGATATTATGCGTGGGTTGGCATCGCTTACCTGTTATGTTTGGCGTATGCGAGC 1284
Db 301 GCTCGATATTATGCGTGGGTTGGCATCGCTTACCTGTTATGTTTGGCGTATGCGAGC 360
QY 1285 GAAGACGCCATGACAAACAAAGGTGAAGCGCCACAGATCATGTAGAGAACAGACCAAC 1344
Db 361 GAAGACGCCATGACAAACAAAGGTGAAGCGCCACAGATCATGTAGAGAACAGACCAAC 420
QY 1345 CGTGGCCGATGACAGCGCTTGGCGGTTCCGCGGTGGCCACTGACAGCGCAACCGGCT 1404
Db 421 CGTGGCCGATGACAGCGCTTGGCGGTTCCGCGGTGGCCACTGACAGCGCAACCGGCT 480
QY 1405 CGCGGTGGAGTGAAGCGTGGATGAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGGCAAT 1464
Db 481 CGCGGTGGAGTGAAGCGTGGATGAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGGCAAT 540
QY 1465 CGTCTGACCTGTTGAACCCGAAATGCTATTTGACGCGCTTGTGTTTATCGCGGCGT 1524
Db 541 CGTCTGACCTGTTGAACCCGAAATGCTATTTGACGCGCTTGTGTTTATCGCGGCGT 600
QY 1525 CGGCGCGCAATAGCGGACACCGAGCGGTGGATTTTCCGCGCTGGCGGCTTGGCGGCAAG 1584
Db 601 CGGCGCGCAATAGCGGACACCGAGCGGTGGATTTTCCGCGCTGGCGGCTTGGCGGCAAG 660
QY 1585 CTTGATCTGGTCCCGCTGGTGGTTCGCGGACAGCATGTGACGCGCGCTGTCAG 1644
Db 661 CTTGATCTGGTCCCGCTGGTGGTTCGCGGACAGCATGTGACGCGCGCTGTCAG 720
QY 1645 CCCCAGGTGCGGCTGGATCAAGCTGCTGCGGAGTTGTGATGACGCGATTGGCCAT 1704
Db 721 CCCCAGGTGCGGCTGGATCAAGCTGCTGCGGAGTTGTGATGACGCGATTGGCCAT 780
QY 1705 CAACCTGATGTTGATGGTTAGTTTTCGCGGTTTGGAAATC 1746
Db 781 CAACCTGATGTTGATGGTTAGTTTTCGCGGTTTGGAAATC 822

RESULT 8
AAS96098 standard; DNA; 822 BP.
AC AAS96098;
AC AAS96098;
DT 26-FEB-2002 (first entry)
XX C. glutamicum gene #23 encoding metabolic pathway protein.
DE Metabolic pathway protein; MP; lysine biosynthesis pathway;
KW methionine biosynthesis pathway; large-scale production of fine chemical;
KW Corynebacterium diphtheriae; diphtheria; ds.
XX Corynebacterium glutamicum.
XX WO200166573-A2.
XX 13-SEP-2001.
XX 22-DEC-2000; 2000WO-IB002035.
XX 09-MAR-2000; 2000US-0187970P.
XX 23-JUN-2000; 2000US-00606740.
XX (BADI) BASF AG.
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberkauer G;
XX Kim J, Lee H, Hwang B;
XX

DR WPI: 2001-582269/65.
DR P-PSDB; AAU71888.
XX Nucleic acids encoding metabolic pathway proteins from Corynebacterium
PT glutamicum, useful for producing methionine and lysine in Corynebacterium
PT and Brevibacterium.
XX Disclosure; Page 214-215; 316pp; English.
XX The present invention relates to the isolation of novel Corynebacterium
CC glutamicum genes encoding metabolic pathway (MP) proteins (AAU71883-
CC AAU71922). The metabolic pathway proteins of the invention include
CC enzymes involved in the lysine and methionine biosynthetic pathways. The
CC polynucleotide sequences of the invention can be used for the large-scale
CC production and/or modulation of expression of fine chemicals such as
CC lysine and methionine. The sequences of the invention may be used to
CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a
CC subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum
CC genes encoding the novel metabolic pathway proteins of the invention
XX
XX Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;
SQ
Query Match 34.6%; Score 822; DB 4; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.2e-245;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 925 GAGCAAGTGTCCAGTGTGAATGGGGTTCATGAAGCTATATTAACCATGTTAAGAACCA 984
Db 1 GAGCAAGTGTCCAGTGTGAATGGGGTTCATGAAGCTATATTAACCATGTTAAGAACCA 60
QY 985 TCATTTTACTTAAGTGTTCATAGTTCATAGTTCATAGTTCATAGTTCATAGTTCATAGT 1044
Db 61 TCATTTTACTTAAGTGTTCATAGTTCATAGTTCATAGTTCATAGTTCATAGTTCATAGT 120
QY 1045 TCTGCTTTTGGGGCCAGTCTTTTACTGTCCATGCGGCGAGATGTACTGTGATTA 1104
Db 121 TCTGCTTTTGGGGCCAGTCTTTTACTGTCCATGCGGCGAGATGTACTGTGATTA 180
QY 1105 ACAAGAAATTAAGCGCGAGGACTCATTTGGGTTCTTCTCGTGTTTAAATTTCTGAGCT 1164
Db 181 ACAAGAAATTAAGCGCGAGGACTCATTTGGGTTCTTCTCGTGTTTAAATTTCTGAGCT 240
QY 1165 CTTTGTGTTCAATCGCGGCACTTGGGGTTGATCTTTTGTCCAAATCGCGCGGATCGT 1224
Db 241 CTTTGTGTTCAATCGCGGCACTTGGGGTTGATCTTTTGTCCAAATCGCGCGGATCGT 300
QY 1225 GTCGATATTATGCGCTGGGGTGGCATCGCTTACCTGTTATGTTTGGCGTATGCGAGC 1284
Db 301 GTCGATATTATGCGCTGGGGTGGCATCGCTTACCTGTTATGTTTGGCGTATGCGAGC 360
QY 1285 GAAGACGCCATGACAAACAAAGGTGAAGCGCCACAGATCATTTGAAGAACAGAACCAAC 1344
Db 361 GAAGACGCCATGACAAACAAAGGTGAAGCGCCACAGATCATTTGAAGAACAGAACCAAC 420
QY 1345 CTTGCGCGATGACAGCGCTTGGCGGTTGGGTTAAAGCCCATGTTGATGGCAAT 1464
Db 421 CTTGCGCGATGACAGCGCTTGGCGGTTGGGTTAAAGCCCATGTTGATGGCAAT 540
QY 1465 CTTGCGCGATGACAGCGCTTGGCGGTTGGGTTAAAGCCCATGTTGATGGCAAT 1524
Db 541 CTTGCGCGATGACAGCGCTTGGCGGTTGGGTTAAAGCCCATGTTGATGGCAAT 600
QY 1525 CGGCGCGCAATAGCGGACACCGAGCGGTGGATTTTCCGCGCTGGCGGCTTGGCGGCAAG 1584
Db 601 CGGCGCGCAATAGCGGACACCGAGCGGTGGATTTTCCGCGCTGGCGGCTTGGCGGCAAG 660
QY 1585 CTTGATCTGGTCCCGCTGGTGGTTCGCGGACAGCATGTGACGCGCGCTGTCAG 1644
Db 661 CTTGATCTGGTCCCGCTGGTGGTTCGCGGACAGCATGTGACGCGCGCTGTCAG 720

Thu Mar 18 12:30:59 2004

1676 GTGGCAGTGTGATGACCGCATTTGCCATCAAACTGATGTTGATGGGT 1723
 661 GTGGCAGTGTGATGACCGCATTTGCCATCAAACTGATGTTGATGGGT 708

QY
 Db

RESULT 11
 AAH45375

ID AAH45375 standard; DNA; 1568 BP.

XX

XX AAH45375;

XX

XX 11-SEP-2001 (first entry)

XX

XX C. thermaminogenes lysin biosynthetic enzyme lyse DNA.

XX

XX Heat-resistant; lysin biosynthesis; enzyme; coryneform;

XX

XX aspartate-semialdehyde dehydrogenase; lyse; ds.

XX

XX Corynebacterium thermoaminogenes.

XX

XX JP2001120270-A.

XX

XX 08-MAY-2001.

XX

XX 01-NOV-1999; 99JP-00311148.

XX

XX 01-NOV-1999; 99JP-00311148.

XX

XX (AJIN) AJINOMOTO KK.

XX

XX WPI; 2001-364760/38.

XX

XX P-PSDB; AAG64047.

XX

XX A heat-resistant lysin biosynthetic system enzyme gene of a high

XX

XX temperature-resistant coryneform microbe.

XX

XX Example 5; Page 22-24; 27pp; Japanese.

XX

XX The invention relates to a gene from a high temperature-resistant

XX

XX coryneform microbe that encodes a heat-resistant lysin biosynthetic

XX

XX enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and

XX

XX can be used for growing amino acid-producing microbes. The present

XX

XX sequence encodes an enzyme of the invention

XX

XX Sequence 1568 BP; 258 A; 525 C; 490 G; 295 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WPI; 2001-376931/40.
 DR P-PSDB; AAG93201.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutanicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids.
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WPI; 2001-376931/40.
 DR P-PSDB; AAG93201.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutanicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids.
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WPI; 2001-376931/40.
 DR P-PSDB; AAG93201.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutanicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids.
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WPI; 2001-376931/40.
 DR P-PSDB; AAG93201.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutanicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids.
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY
1679 GCAGTTGTGATGACCGGCATTGGCCATCAAACGATGTTGATGGTTA 1725
1726 GCGCTGTCCTTCCCGCATTTGGCCTGAAGTGATCCTGATGGTTA 1772

RESULT 12	
ACC80942	
ID	ACC80942 standard; DNA; 712 BP.
XX	
XX	ACC80942;
XX	
XX	27-OCT-2003 (revised)
DT	DT
DT	11-AUG-2003 (first entry)
XX	
XX	
XX	LysE24 protein encoding sequence.
DE	
XX	
XX	L-lysine; L-arginine; LysE24; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
Key	Location/Qualifiers
PH	1. .711
FT	/tag= a
CDS	/product= "lysE protein"
FT	
FT	
FT	

RP1266966-A2.

18-DEC-2002.

05-JTN-2002: 2002EP-00012539.

12-TTN-2001: 2001JP-00177075.

(A.TTN) ATTNOMOTO CO INC.

amiti v vaspreda H.

0003-2467

P-PSDB; ABR58214.

Novel DNA encoding va

excretion of L-lysine

Example 1; Page 19-20; 23pp; English.

The present invention relates to DNA encoding variants of protein with a loop region and six hydrophobic helices which facilitates excretion of L-lysine and/or L-arginine to outside of cell of a methanol assimilating bacterium when introduced into the bacterium. The method is used for encoding a protein which facilitates excretion of L-lysine, L-arginine or both of these L-amino acids to outside of a cell of a methanol assimilating bacterium when DNA of the method is introduced into the bacterium. The present sequence represents a lyx24 protein from *Brevibacterium lactofermentum* encoding sequence. (Updated on 27-OCT-2003 CC BY-NC-ND 4.0 International license. <https://doi.org/10.1101/000000>)

Sequence 712 BP: 133 A; 173 C; 224 G; 182 T; 0 U; 0 Other;

29.4%; Score 696.8; DB 8; Length 712;

1st Local Similarity 99.6%; Pred. NO. 1.00e+000
0: Mismatches 2; Indels 1

1016 ATGCTGTCATCGGAATCTTTTCACTTACAGGTCGTGCTTTGGGGGCCAGTCTTTT

1 ATGCGGCGCTGCGGATATGCTCAATTAACAGGCTGCTGCTTTTGGGGGCCAGTCTTTT

...TGGCGGCGCCGCAATCTATCTGCTGATTAAACAAGGAATTAGCGCGAAGGAC

ATGCGCAGTCTCATTAACAAGGAATTACGGCGAAGGAC

CCCTGTTTCTTCATCGGGGCACCT

599	GTGACGCTGTCTCCACCCCAAGAGACTACTCTGTTGAACACGGGAGGAAACATAGTGGAT	858	QY
361	GTGAGGGTGAACCGGACCCCAATGTGTGAACCTCGGCGACACGGGCGGGAACCAAGTGGAC	420	DB
659	AGCGAATCTGCGTTGATGGCGATGTTAAACGGGATTTCAAGAGGCGTCAGATAGTGC	718	QY
421	AGGGAATGGCGTTGATAGCCACGCTCAGGGGATCTCGTCAAGCCGTTTCGCCAGTTGC	480	DB
719	GCTTTAGTTTCTGCTTGGAGGACACCACTTTTCCGCGTCTTGTGCAAAAGACTTCACCC	778	QY
481	TCAACGGGTCTCCGCTGCGAGAGCGCCATCTTCGCGGCGCTGGACGACACCTCCCG	540	DB
779	GCTTCGGTTGCTTTTGGCCGCTTGGGTGGCGGATACCAACTCGACCCAGTGATGCTCG	838	QY
541	GCTTCGGTGGCAACCGCGCGTGGGTGGCGGACCAACAGTAGCCGACCGATTTCTCC	600	DB
839	AGAGCTTTAAACGCGCTGACTACCGCGAGGGGGAATGGAAGGCTAAGGAGCGGCT	898	QY
601	AGTGCCCTTGATGCGCTCGGCTGCCCGCCGAGGGGGAATCGACGTGCGAGGAGCGTTC	660	DB
899	TCCAGCTGCGCTTCATCAATGATTCAGAGCAAAAGTCTCCAGTTGAATGGGGTTCAATGAAG	958	QY
661	TCGAGCTGCGCTCTGTCGATGATGTCAGAGGGTGTCCAGGTGATCGGGTTTCATGAAG	720	DB
959	CTATATTAAACCATGTTAAGAACCAATCATTTTACTTAAGTACTTCCATAGTTCACGATG	1018	QY
721	CAATTTAAACCATGTTCACTCTCTCATCAITTTACTTAAATGTGGTTTGTCCGCGAACATG	780	DB
1019	GTGATCATGGAATCTTCATTACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTGCCATC	1078	QY
781	CGGAGCATGGAATCTTTGTACCGGTTTGTGTGGAGCGACGTCGCTGTTTGGCCATC	840	DB
1079	GCACCGCAGATGTACTGGTGATTAAACAGGAATTAAGCGCGAAGAGACTCATTTGGGTT	1138	QY
841	GGCCACAGAGATGCTCTGGTGATCAACAGAGGACATCAACCGGAGGGCATCAGCGCGCTC	900	DB
1139	TTTCTCGTGTGTTAAATTTCTGACGCTCTTTTGTTCATCGCCGCACTTTGGCGCTGAT	1198	QY
901	ATCATCTGCTCTGCTGTCGCGAGTGGTCTGTTTACCTTCGGCACCTTCGGGCTCGGC	960	DB
1199	CTTTTGTGCAATGGCGCGCGATCGTCTCGATATTATGCGCTTGGGGTGGCATCGCTTAC	1258	QY
961	CTGATCTCGACACCGCCCCGATCATCTTCGACATCTCGGCTGTGTGGCGCATCGCCTAC	1020	DB
1259	CTGTTATGTTTGGCGTCAITGGCAGCGAAGACGCGATGACAAAACAGGTGGAGCGCA	1318	QY
1021	CTGCTGTGTTTGGCGTGATGGCGGCGCGACGCGCTTGGCGCCGACCGAGGTAAAC	1080	DB
1319	CAGATCATTTGAAGAACAGAACCAACCGTGCCCGATGACACGCTTTGGCGCGTTCCGGG	1378	QY
1081	TTTGTGCT-----GAGCATTCGGAACCCGTTG---CGCAGCGTTCGCTCCGCG	1125	DB
1379	GTGGCCATCGACCGCGAAACCGGTTGCGGGTGGAGGTGAGCGTTCGATAAGACGCGGTT	1438	QY
1126	GGGGCGGTGACGACGAAACGACACCGCCCGCTCCGCATCATCAGGCAACCGGAGGTC	1185	DB
1439	TGGGTAAAGCCGATGTTGATGGCAATCGTGCTGACCTGGTTGAACCCGAATGCGTATTG	1498	QY
1186	TGGGTCAAGGCCCATGTCTATGGCCATTTGTGTGACTTGCTGGCTCAATCCCAATGCCCTAC	1245	DB
1499	GACGCGTTTGTGTTTATCGGCGCGGTTCGGCGCGCAATACGGCGACACCGGACCGTGGAT	1558	QY
1246	GATGCTCTGCTCTTCATTCGGTGTGTTCGAGGCCAGCTACGGGGAGACCGTCCGTTGATC	1305	DB
1559	TTCCGCGCTTGGCGGTTTCGCGCAAGCGCTGATCTGTGTTCCGCTGTGTGGGTTTCGGCGCA	1618	QY
1306	TTGCTGCGGGTGCTTTCCGCCCAAGCTGTGTGTTTCCCATCTGCTGGTTTCCCATCTGGTTC	1365	DB
1619	GCAGCATGTTCACGCCCGCTGTTCAGGCCCAAGGTGTGGCGCTGCTGATCAACCTCGTGTG	1678	QY
1366	GCAGCATGTGTGCGCTGCCCTGTGTCCTCCCGCGGGCTGTGGCGCTGATCAACATAGTGTG	1425	DB

Db	121	GTCTCTCGGTGTGTTAAATTCTGACGCTCTTTTGTGTTCAATCCGCGCACCTTGGGCGTT	180
Qy	1196	GATCTTTTGTCCAAATGCCGCGCCGATTCGTCTCGATATATTGCTGCGGTGGGATGCATCCT	1255
Db	181	GATCTTTTGTCCAAATGCCGCGCCGATTCGTCTCGATATATTATGCGCTGGGATGCATCCT	240
Qy	1256	TACCTGTTATATGTTTGGCGTTCATGCACGAGAAAGACGATGACAAACAAGGTGGAAAGCG	1315
Db	241	TACCTGTTATATGTTTGGCGTTCATGCACGAGAAAGACGCGATGACAAACAAGGTGGAAAGCG	300
Qy	1316	CCACAGATCATTCAGAAACAGAAACCAACCTGTCGCCGATGACAGCGCTTTGGGCGCGTTC	1374
Db	301	CCACAGATCATTCAGAAACAGAAACCAACCTGTCGCCGATGACAGCGCTTTGGGCGGTTC	360
Qy	1375	GCGCGTGGCCACTGACACGCGCAACCGGGTGGGGTGGAGTGGAGCTGCATATAGCACGCG	1434
Db	361	GCGCGTGGCCACTGACACGCGCAACCGGGTGGGGTGGAGTGGAGCTGCATATAGCACGCG	420
Qy	1435	GGTTTGGGTAAAGCCCATGTTGATGAGCAATCGTGTGCTGATCGTTGTAACCCGAATGCCTA	1494
Db	421	GGTTTGGGTAAAGCCCATGTTGATGAGCAATCGTGTGCTGATCGTTGTAACCCGAATGCCTA	480
Qy	1495	TTTGACGCGCTTGTGTGTTATCCGGCGCGTCCGCGCGCAATACGCGCGACACCGGACGGTG	1554
Db	481	TTTGACGCGCTTGTGTGTTATCCGGCGCGTCCGCGCGCAATACGCGCGACACCGGACGGTG	540
Qy	1555	GATTTTCGCGCTGTGGCGGTTTCGCGGCAAGCCCTGATCTGTGTTCCGCTGTGTGGGTTTCGG	1614
Db	541	GATTTTCGCGCTGTGGCGGTTTCGCGGCAAGCCCTGATCTGTGTTCCGCTGTGTGGGTTTCGG	600
Qy	1615	CGCAGCAGCATTTGTCAACCCCGCTGTCCAGAGCCCAAGTGTGGCGCTGGATCAAGCTGCT	1674
Db	601	CGCAGCAGCATTTGTCAACCCCGCTGTCCAGAGCCCAAGTGTGGCGCTGGATCAAGCTGCT	660
Qy	1675	CGTGGCAGTTGTGATGACCGCATTTGCCCATCAAACTGATGTTGATGGGTTAG	1726
Db	661	CGTGGCAGTTGTGATGACCGCATTTGCCCATCAAACTGATGTTGATGGGTTAG	712

RESULT 13

ID	AAH68419/C
XX	AAH68419 standard; DNA; 627 BP.
XX	AAH68419;
GC	
XX	
DT	26-SEP-2001 (first entry)
XX	
DE	C glutamicum coding sequence fragment SEQ ID NO: 3454.
XX	
KW	Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW	organic acid synthesis; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
PN	EP1108790-A2.
XX	
PD	20-JUN-2001.
PF	
XX	18-DEC-2000; 2000EP-00127688.
XX	
PR	16-DEC-1999; 99JP-00377484.
PR	07-APR-2000; 2000JP-00159162.
PR	03-AUG-2000; 2000JP-00280388.
XX	
PA	(KYOW) KYOWA HAKKO KOGYO KK.
PI	Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI	Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX	
DR	WFI: 2001-376931/40.
XX	P-FSDB; AAG93200.
PT	
XX	Novel polynucleotides derived from Coryneform bacteria, for identifying

Thu Mar 18 12:30:59 2004

1970 TCCTGGCGATGTCAATGAGTTCGGGACCATATCAATATTTGTTTCAGTTTCAACATGCC 2029
 854 CGTAAACGCTAAATCAATTCAGTGCACGACCATGTCAGATTTTAGGTTGAGCATCTCT 795
 2030 TCAGACAGGACTTACCTGGTGGCGGGAACCTCTTGAATTCATCGAGATATTG 2089
 794 TTGGCCAAAGGACTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 735
 2090 TCCGTGACGAGCCCTCGCGCAAGTGGTGAAGAAATGACGCCAAGACCATTTGTCGA 2149
 734 TCAGTTAAAGCCCTCTGCAAGTGGCGAAATGCAATAACACCAAGTCCGTTATTTCGCC 675
 2150 GTGACTGGAACAAGTTCTCAGCTCATCGCCGCTTCTCCACCCCAACGATTAAATGATG 2209
 674 GCAGACTCTAAACAGTTCTCGCCATCTTCACCTGGTTCTTCTACCCAGCGTTCAAGATA 615
 2210 GAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCCGCTATGAACTCAGCGGCC 2269
 614 GAATAGCTCGGTTGATGATCAGAGCGGCGAGCCCTCACCAGCTAGAATTTCCGAGATC 555
 2270 TCCGCTGTGAGCTCTGGACCGTGAAGAAATACCCAGTAAAGAGCCCTTCCAGACGCA 2329
 554 TCTGCGGTTAGTTCAAGTTCATAGGAGGAAATTCACATAGCGAGCCCTTCCCGGATGCC 495
 2330 ACAATGTCAGCAATGCGTACATGTTCTTCCAAAGAGTATCT 2374
 494 ACGATGTCTCGCAAGACATATGCGAGTTTCTTCCAAAGCGGTATCT 450

Search completed: March 15, 2004, 14:59:10
 Job time : 974.046 secs

PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR P-PSDB: ABU25781.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 14; SEQ ID NO 17521; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway; (8)
 CC required for proliferation, or that inhibits cellular proliferation; (9)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC not form part of the printed specification. Note: the sequence data for this patent did
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1095 BP; 264 A; 247 C; 311 G; 273 T; 0 U; 0 Other;
 SQ
 Query Match 12.9%; Score 305.8; DB 7; Length 1095;
 Best Local Similarity 67.1%; Pred. No. 3.5e-84;
 Matches 433; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
 1730 TCGGGGTTTGGATCGGTGGCTTCGCCCAATGTTGATGCGCGGTCTGGGAAATC 1789
 1094 TCGTGAACCTTCGAGCGAGTAGCACCCGCCCAATATTAATTCAGCGTCTGGCAACA 1035
 1790 TCATGATCGCTCCCACTCGGCTCAGAAATCCCAAGTGTGATGATCAGGCTG 1849
 1034 GTGTCAATATATGCGTTCTTCAACGCTGAACATTAATAGAGCTCCAGATTT 975
 1850 TTGTCCAGCTGTCACTGACGAGACCAATCAATGCTGCTACGCTATCCGCGCG 1909
 974 TGATCAACATGCGCTACCGAGATGCTCCGATCAACGACCTGGTTACCGTTGGCCCCA 915
 1910 TACTCTCTTGTTCGCGAGCACCCATGCAAGCGCATCTGCGCAAGTCACTGCGCGGT 1969
 914 TAATCGCTTGTTCGCGAGCACCCAGCAATAGCAGCTGTGCAACGTTTGGCCAGT 855

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:47:35 ; Search time 178.837 Seconds
(without alignments)
7366.769 Million cell updates/sec

Title: US-09-105-117K-1
Perfect score: 2374
Sequence: 1 ccatttctgaagggtttac.....tttttccaaaggatattct 2374

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Parents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
C 1	122.8	5.2	15239	1	US-08-390-878-17	Sequence 17, Appl
C 2	121.2	5.1	909	4	US-09-894-844-12	Sequence 12, Appl
C 3	121.2	5.1	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 4	121.2	5.1	4411529	3	US-09-103-840A-1	Sequence 1, Appl
C 5	100.8	4.2	5541	1	US-08-920-812-20	Sequence 20, Appl
C 6	100.8	4.2	5541	1	US-08-920-827-20	Sequence 20, Appl
C 7	100.8	4.2	5541	1	US-08-921-177-20	Sequence 20, Appl
C 8	100.8	4.2	5541	1	US-08-362-577C-20	Sequence 20, Appl
C 9	100.8	4.2	5541	2	US-08-920-828-20	Sequence 20, Appl
C 10	81.4	3.4	1095	4	US-09-489-039A-5370	Sequence 5370, Ap
C 11	80.2	3.4	915	4	US-09-489-039A-4735	Sequence 4735, Ap
C 12	73.4	3.1	936	4	US-09-543-681A-791	Sequence 791, Ap
C 13	65.2	2.7	900	4	US-09-252-991A-7644	Sequence 7644, Ap
C 14	64	2.7	1002	4	US-09-724-623-24	Sequence 24, Appl
C 15	61.8	2.6	699	4	US-09-252-991A-7908	Sequence 7908, Ap
C 16	61.8	2.6	834	4	US-09-252-991A-7643	Sequence 7643, Ap
C 17	61.8	2.6	894	4	US-09-252-991A-7841	Sequence 7841, Ap
C 18	57	2.4	597	4	US-09-894-844-11	Sequence 11, Appl
C 19	56.6	2.4	642	4	US-09-489-039A-4674	Sequence 4674, Ap
C 20	54.6	2.3	903	4	US-09-328-352-1694	Sequence 1694, Ap
C 21	53.8	2.3	1176	4	US-09-252-991A-7571	Sequence 7571, Ap
C 22	49.2	2.1	921	4	US-09-252-991A-14393	Sequence 14393, A
C 23	49.2	2.1	1905	4	US-09-252-991A-14312	Sequence 14312, A
C 24	47.8	2.0	474	4	US-09-252-991A-7572	Sequence 7572, Ap
C 25	46.2	1.9	903	4	US-09-252-991A-5488	Sequence 5488, Ap
C 26	45.8	1.9	720	4	US-09-252-991A-7723	Sequence 7723, Ap
C 27	45.4	1.9	45613	4	US-09-596-002-22	Sequence 22, Appl

Sequence 800, App
Sequence 142, App
Sequence 16286, A
Sequence 15990, A
Sequence 16513, A
Sequence 15639, A
Sequence 345, App
Sequence 5423, Ap
Sequence 12026, A
Sequence 11957, A
Sequence 9, Appli
Sequence 15639, A
Sequence 12324, A
Sequence 12096, A
Sequence 12421, A
Sequence 12055, A
Sequence 2, Appli
Sequence 1, Appli

28 43 1.8 645 4 US-09-543-681A-800
C 29 42.8 1.8 1879 4 US-09-453-702B-142
C 30 42 1.8 936 4 US-09-252-991A-16286
C 31 42 1.8 978 4 US-09-252-991A-15990
C 32 42 1.8 1533 4 US-09-252-991A-16513
C 33 41.6 1.8 505 4 US-09-621-976-15639
C 34 40.6 1.7 969 4 US-09-252-991A-5423
C 35 40.4 1.7 285 4 US-09-252-991A-12026
C 36 40 1.7 864 4 US-09-252-991A-11957
C 37 40 1.7 870 4 US-09-492-027A-9
C 38 39.6 1.7 1935 2 US-08-621-976-15639
C 39 39.4 1.7 505 4 US-09-252-991A-13324
C 40 39.4 1.7 927 4 US-09-252-991A-12096
C 41 39.4 1.7 972 4 US-09-252-991A-12096
C 42 39.4 1.7 1482 4 US-09-252-991A-12421
C 43 38.6 1.6 924 4 US-09-252-991A-12055
C 44 37.6 1.6 4403765 3 US-09-103-840A-2
C 45 37.6 1.6 4411529 3 US-09-103-840A-1

ALIGNMENTS

RESULT 1
US-08-390-878-17/c
; Sequence 17, Application US/08390878
; Patent No. 5700683
; GENERAL INFORMATION:
; APPLICANT: Stover, Charles K.
; APPLICANT: Mahairas, Gregory G.
; TITLE OF INVENTION: VIRULENCE-ATTENUATING GENETIC DELETIONS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower, 20th
; STREET: Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,878
; FILING DATE: 17-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 15371A-17
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/543/9600
; TELEFAX: 415/543/5043
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15239 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-390-878-17

Query Match 5.2%; Score 122.8; DB 1; Length 15239;
Best Local Similarity 49.7%; Pred. No. 1.1e-25;
Matches 434; Conservative 0; Mismatches 422; Indels 18; Gaps 4;
QY 78 GTAACCTAAGCGCGCAATTCCTGATTCGATCATCAGACGCGCTCTGTGAGTCTAGCTA 137
DB 6117 GTTGTGCGCGCGGTACAGACGCGTTGCGCGCGCCCTACCGGTGTCGGTAATTCGCGCA 6058

; PRIOR APPLICATION NUMBER: 60/097,936
 ;
 ; PRIOR FILING DATE: 1998-08-25
 ;
 ; NUMBER OF SEQ ID NOS: 137
 ;
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ;
 ; SEQ ID NO 12
 ;
 ; LENGTH: 909
 ;
 ; TYPE: DNA
 ;
 ; ORGANISM: Mycobacteria tuberculosis
 ;
 ; 795-09-894-844-12

Query Match	5.1%	Score 121.2;	DB 4;	Length 909;
Best Local Similarity	49.5%	Pred. No. 5.9e-26;		
Matches 433; Conservative	0;	Mismatches 423;	Indels 18;	Gaps 4;

78	GTAACTAGGCGCGCATCCTCTCGAATGCTGTGCATCAAGACGCGGCTGTGTGAGTCTAGCTA	137
892	GTTCCTGGCGCCCGGTACAGACCGCTGCCCGCGCCCTCACCGTGTGCGGTAAATTCGCGCGA	833
138	GAGATCTTAGATTCACAGCGCCCATGTTGGCCATATCATCTGGTGTGTCAATGGGTATCTCAT	197
832	TGATCGGACTGTCTCCAGTTTTCCAGCATTTGCCAATGAGAGGGAGCTGAGGTGTGTGTGCG	773
198	CGAGGAGATCACTTCTCTCTGCTTTTAGCATGGGAGCAGCTTGGGTTTCGGGGAAGATGC	257
772	AGACCCGATGACGAAGATCCATCGGCAAGCGGATGCTGCGACCTTCTCGGGGAACATCG	713
258	CCCAACACAGGCGCTCGCGGAATGCTCTACCCAAACCTTTGCGCGGACGGGACATGGATA	317
712	CCCATCCAGCCGCGCGCGCTGCGCGGTGAAGCCCTCTGTGTGTCGGGACAAAGTCGG	653
318	CGCGCTTGCGCCCCACAGAGACATCGAC---GGGCGCGCTCCAGGTCACTGCTTTTGAAGCA	374
652	TGCGTCTGCTGTATGCGCGGACGAAGGGCTTTAGCCACCAACATGTCCTCGAGCCCATCGT	593
375	CATCTTTTGGGACCGAAGCGTAAGACGGGCATCGCAGCCCAATCTAGTTTCCATCAACCA	434
592	CACGATTCACGCGCAGTACCGGAGCTTTAGCCGCGCGCGCGCAGTGAACCCGCTCGGATA	533
435	TGTAGGCATCCCGCAATGAGGGGTTGCATGCGCAAGTGGCGCATGGTTCCAAGTTCTTA	494
532	GATGGCGCTGGACGAATGGCTCTCGGCACCTGATAGTGTAGGTAGCGCATTTCAACGAGCGGT	473
495	CTTACTTCATCATCCGCGCACGGGATAGCTTCACGGGTTACCGCTCTCTAATAACTCTCCAC	554
472	GCACCCGCGACCCCGGCACCGGGTTTCGCTCGGTGGTCAACCGCGCCCATCGCCACACCTT	413
555	GCGCAGCAGAGGATAATGTGTGGCTTCAATTTCCAAAGCAGCGTGTGCGGTGCTCCAC	614
412	CCCGTAGCAGCGCGCGGAATGGTTCCTGCTCTCGATCCGAACCTCGAGGAGGAGCTCG	353
615	CCGAGAGCTACCTGTTGTAAACCGGAGGAACCATGTGATAGCGAATCTCGGTGA	674
352	-----CGAGACGCTGAAACACGCGCGGAACCAATGTCGCAATCGAATCGCGGTTTA	302
675	TGCGATGGTTAAACGGGA-----TTTCAGCAGAGCGT---CCAGATAGTGGCTTTAGTTT	728
301	CCGCAATGTGATCCGCGTGGCTTTTCAGGCAACGGTGTGCCACCCATTTCAGCGAGCGCT	242
729	CTGCTTGACAGAACCAATTTTCGCGCTGCTGTGCACAGGACTTCACCGCGCTTCGGTTG	788
241	CGACTCGAGCAACGCTGTTTTCGCGCGCCCAACCGCAACAGCGGATACCTCGGTGCTG	182
789	CTTTGGCCGCTTGGTGGCGGATACCAACTCGACCACTGTATGCTCGAGAGCTTTAA	848
181	CCCGACATGGCTTTTCCCTTGACCCACAGACACTGGCCGACCTGCTGCTCCACGACTTGA	122
849	CGCGCTGACTCAACCGCGAGGGGGAATGGAAGGGCTTAGGAGGCGCTTTCGAGCTGC	908
121	TGCGCTGACTGACAGCGGACCGGCGGTGAATGATAGCGGCTTCGCGCGCATCGAGCTGC	62
909	CTTCACTCATATGATTGAGGCAAGTGTCCAGTTG	942
61	CCAGTTTCGACCAACGCGACGCCAATGCGCGCAGCTG	28

138	GAGATCTTAGATTCCAGGCGCCATCGTTGGCCAAATACATCGGTGTCTCAATGGGTATCTCAT	197
6057	TGATCGCAGTGTCCAGTTTCCAGCATTTGCCAATTAGAGGCGACGTGAGGTGTATGTGCG	5998
198	CGAGGAGGATCATCTTCTCTCGTTTATAGCATGGGAGCAGCTTGGGTTTCGGGAGAGATC	257
5997	AGACCCGTGACACGATCCATCGGCAAGCGAGATGCTGCCAGCTTCTCGGGCAACATGC	5938
258	CCCAACCAAGGCTTCGGCGAATTGCCTACCAAAACCTTCCGCGACGGGACAAATGGATA	317
5937	CCCATCCAGCCGCGCGCGCTCGCGCGGTGAAGCCCTCTGTGTCGGGACAAAGTGGC	5878
318	CGCGCTCGGCCCCACAGACCACTCGAC--GGGCGCGGTCCAGGTACAGGTCTTGAGACA	374
5877	TCGGTCTGGTATGGCGCGACGAAGGCCCTTACGCCAACCAATGCTCTGCAGCCCATCGT	5818
375	CATCTTTGGGACCGAAGCGTATAGACGGGCATCCGAGCCCAATCTAGTTTCCCATCAACA	434
5817	CAGATTCCAGCCAGTGAACGAGGCTTTAGCCGCGCGCGCGGAGTGACCCGTCCGATA	5758
435	TGTAGGCATCCCGCAATGAGGGGCTTGCAATGGCCAAAGTGGCGCATGGTTTCCAAGTTCTA	494
5757	GATGGCGCTTGACGAATTGGCTGCTGGGCCACTGGTAGGTAGCGATTTCAACCAGCGGCT	5698
495	CTACTTTCACATCCCGGCCACGGGATAGGTTACGGGTTACCGTCTCTAAACATCTCCAC	554
5697	GCACCGGCGAGCCCGCACCGGTTCCGCTCGGTGCTCACCGCGCCCATCGCCACACCTT	5638
555	CGCGTAGAAGGATTAATGTGTGGCTTATCTTCCAAAGCGACGCTGAGAGGTGTGCTCCAC	614
5637	CCCGTAGACGCGCGCGAATGTTCTGTGTTCTCGATCCGACGTCCAGCAGGACGTGCG	5578
615	CCCAAGAAGCTACTCTGTTGAAACCGGGAGGAAACCAATGAGATACGAATCTCGCTGA	674
5577	-----CGAGACGCTGGAACGCGCGGAACCATGTGCGCATGGAATCGCGTTTA	5527
675	TGGCGATGGTTTAACGGGA-----TTTCAGCAAGCGT--CCAGATAGTTGCGCTTTAGTTT	728
5526	CCGCAATGGTGTATCCGCGTGGGTTTCAGCGACGGGTTGCCACCCTTTTCAGCGAGCGCT	5467
729	CTGCTTGCACACACACCAATTTTCGGGCTGTTTGCACAAAGGACTTCAACCGCTTCGGGTG	788
5466	CGGACTCGAGCAACGTTGTTTCGGGCGCAACCGGCAACAGCGGGATACCTGCGGCTCTCG	5407
789	CTTTGGCGGGTTGGGTGGCGGATACCAACACTGACCCACCGTGAATCTCGAGAGCTTAA	848
5406	CCGACATGGCTTTTCCCTGACCAACGACACTGGCGGACCTGCTGCTCCAAAGCACTTGA	5347
849	CGCGCTGACTACCGCGGAGGGGNAATGAAGAGGCTTAAGGAGGGCGCTTTCGAAGCTGC	908
5346	TGCGCTGACTGACAGCGGAGGGGGTGACATGTAGGCGCTCCGCGCGGCATCGAAGCTGC	5287
909	CTTTCATCAATGATTGAGAGCAAGGTGTCCAGTTG	942
5286	CGACTTCACACCGGACCAATGCGGCAAGCTG	5253

```

RESULT 2
US-09-894-844-12/c
; Sequence 12, Application US/09894844
; Patent No. 6886166
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Smal, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; TITLE OF INVENTION: the M. Tuberculosis Complex
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25

```

675 TGGCGATGTTAACGGA-----TTTCAGCAAGCGT--CCAGATAGTTGGCGCTTAGTTT 128

Thu Mar 18 12:30:59 2004

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 5541 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
STRAIN: Clinical Isolate EC-625
US-08-920-812-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;
Best Local Similarity 50.5%; Pred. No. 2.2e-19;
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps

QY	1623	CATTGTACCGCCGCTGTCAGCCCAAGGTGGCGCTGGATCAACGTCGTCGTCGGAG	1682
Db	2687	CATTATCCATCTGACGTTTGTGCTCTCTGCAATTTGTGATGGCTATCTTACCGACTGAC	2746
QY	1683	TTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGTGTAGTTTTCGGGTTTGG	1742
Db	2747	TGCTCATTAATCACTGATCTTATGCGGCGCGTTCCTGCGCCCGTTATTAAACAGGTCAAT	2806
QY	1743	AATCGTGGCTTCGCCCAAAATTTGATGCGCGCTGTCGCGGAAATCTCATCGATCGCT	1802
Db	2807	TATCGGAAGACGCTGCCACAGATTTCAGCTCGCCATCGCCATCGGCGATATGCTGATCAATCTGG	2866
QY	1803	CGACTCGGCTGAGAAATCTCAAGTTTGTGATGTCGCGCTGTCGCGGAAATCTCATCGATCGCT	1862
Db	2867	CCAGTCTCTCGTGTCTTAATGTGATGATTAATTCAGCGCTGTCGCGGAAATCTCATCGATCGCT	2926
QY	1863	CAACTGACGAAAGCACCAATCAATGCACTGTCGCGCTGTCGCGGAAATCTCATCGATCGCT	1922
Db	2927	CGCGCGGCTGCGCAATCAATACCGACGTCAGCGGATCA-----T 2967	
QY	1923	CGCGCAGACCGCTGCGCGCAATCTGCGCAAGTGAATGTCGCGCTGTCGCGGAAATCTCATCGATCGCT	1982
Db	2968	CTTTTCAGCAACCGCTTACGCGCTTTCGCGCAATGTCGCGCTGTCGCGGAAATCTCATCGATCGCT	3027
QY	1983	CATTGAGCTTGGCGCAATCAATATTTTTCAGTTTCAATGTCGCGCTGTCGCGGAAATCTCATCGATCGCT	2042
Db	3028	CATTCAATTAAGTGTAGCTGTTGAGTTGGCTTGGTTCGTTAGCATTTTTCGCGCTGTCAGACCAAC	3087
QY	2043	TACCTGCGCTGGC-----GCGGGAACCTCTGGAATTCATTCATCGATGATTTT	2090
Db	3088	GACTTATTTTCCCTTCAGATGATCCGTTGATTTTCGCGGAAATGCGCTGAGATATTTTC 3147	
QY	2091	CCGTGAGCAGGCTTGGCGCAATGTCGCGCAAGTGAATGTCGCGCTGTCGCGGAAATCTCATCGATCGCT	2150
Db	3148	CGGTGAGCAATCCCTGAGCCAGGAGTGAAGGCATATACAGCCACCGCGTTATTTTGA 3207	
QY	2151	CTGACTGCAACAGTTTCTCACGCTCATCGCGCTGTCGCGGAAATCTCATCGATGATTTT	2210
Db	3208	GGGTATCCAGCA-----GGCGCTTTTATCCACCGCGTTTCAGTAAAT 3252	
QY	2211	AATAGCTTGGCTGATGATGAGAGCGGCGACGCTTCTCGCATGAATCACTCAGCGCCT	2270
Db	3253	TGTACGAAGTTTGTATGAATTAACGCGGAAATTTTCCATCGCGCAGCACTCAACCAATTT 3312	
QY	2271	CGCTGTGAGCTTGGCGCGCTGAGCAATATACCAATGTCGCGCTGTCGCGGAAATCTCATCGATCGCT	2330
Db	3313	TTTGGCTCGCTCTGGCGAGTGAAGGAGTCCCGACATTAAGCGCTTACCGCTTGTGA 3372	
QY	2331	CAATGTACGCAATGCTATCATGTTTCTTCCAAAGGAGTATCT 2374	
Db	3373	CGCATGAGCCAGCGCAGAGCGGTTTCTTCCATCGCGCTATTT 3416	

RESULT 6

US-08-920-827-20
; Sequence 20, Application US/08920827
; Patent No. 5770375

Db	2229369	GATGGCGCTGGACGATGCTGCTGGCCACTGCTGCTAGTAGGCAATTTACCCAGCGGT	2229428
QY	495	CTACTTCAATCCCGCAGCGGTTAGTCTCACGGGTTACCGCTCTCTAAACATCTCCAC	554
Db	2229429	GCACCGGAGCGCGCAGCGGTTCCGCTGGTGTGTCACCGCGCCATCGCCACACCT	2229488
QY	555	GCAGCAGCAGGATTAATGTGTGGCTTTCATCTTCCAAAGCGAGCGGTGCTGCTCCAC	614
Db	2229489	CCCTGAGCAGCGCGGCAATGCTCTGCTCTGATCCGAACTCGAGCAGCAGCTCGC	2229548
QY	615	CCCAAGAGTCACTCTGTTGAACACCGGAGGAAACCATGTGATGCAATCTGCGTTGA	674
Db	2229549	-----CGAGCAGCTGCAACACACGCGCGGAAACCATGTGCGCAATCTGCGCTTGA	2229599
QY	675	TGGCGATGTTTAAACGGA-----TTTCAGCAAGCGCT--CCAGATAGTTTGGCTTTAGTTT	728
Db	2229600	CCGCAATGGTGTATCGCGCTGCTTTGACGACGCTGTCGCAACCCATTTTCAGCGAGCGCT	2229659
QY	729	CTGCTTGCAGCAACACCAATTTTCGCGCTGCTTGCACAGGACTTCACCGCTTTCGTTG	788
Db	2229660	CGGACTCGAGCAGCTGTTTTCGCGCGCCAAACCGCAACAGCGGAGTACCTGCGTCTGCTG	2229719
QY	789	CTTTGGCGGTTGGTGGCGGATACCACTGACCACTGACCGGAGTGTGCTCGAGAGCTTAA	848
Db	2229720	CCCGACATGCTTTTCCCTGACCAACCGAGCCTTGGCGGAGCTGCTGCTTCAACGACTTGA	2229779
QY	849	CGCGCTGACTCACCGCGGAGGCGGAAATGGAAGGCGCTAAGGAGCGCTTTCGAAGCTGC	908
Db	2229780	TGCGCTGACTGACAGCGGAGGCGGATGACATGATGAGGCTGCGCGCGCGCATCGAAGCTGC	2229839
QY	909	CTTTCATCAATGATGAGCAAAAGTGTCCAGTTG 942	
Db	2229840	CCAGTTGACCGCGGAGCGCAATGCGGCGAGCTG 2229873	

RESULT 5

US-08-920-812-20
; Sequence 20, Application US/08920812
; Patent No. 5763188

GENERAL INFORMATION:

APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago

STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,812
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

Thu Mar 18 12:30:59 2004

```
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,827
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Heien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 5541 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
STRAIN: Clinical Isolate EC-625
US-08-920-827-20
Query Match 4.2%; Score 100.8; DB 1; Length 5541;
Best Local Similarity 50.5%; Pred. No. 2.2e-19;
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;
QY 1623 CATTGTCACGCCGCTGTCAGCCCAAGGTGTGGCGTGGATCAACGTCGCTGGCAG 1682
DB 2687 CATTATCCATCTGACGCTTGTGTCTCTCGCATTTGTGATGGGCTATCTTGACCGCTGAC 2746
QY 1683 TTGTGATGACCGCATGGCCATCAAACTGATGTTGATGGGTAGTTTTGGCGGGTTTGG 1742
DB 2747 TCGTCATATCACTGATCTTATGGCGCGGTTCTCGCGCGGTATTACAGTCAIT 2806
QY 1743 AATCGTGGCTTGGCCCAATTTGATGCGCGGCTGTTGGGAAATCTCATCGATCGCT 1802
DB 2807 TATGGGAAGACGCTGCCACAGATTCAGCTGCCCATCGGCATATGCTGATCAATCTGCG 2866
QY 1803 CCAACTCGGCGTCAGAAATCTCAAGTTGTGATGATCAAGGCTGTGTCAGCTGCT 1862
DB 2867 CCAGCTCTCGGTGCTAAATGTGATTTATTCAGCGCTGACGCTCTCTCAAGTTGTC 2926
QY 1863 CAACTGACGAGACCAATCAATGCTGATCGGCTATCGCGCGGTACTCTCTGCT 1922
DB 2927 CGCGCGGC-TGGCACAATCAATACCGACGTCACGGATCA-----T 2967
QY 1923 CGCGCAGCACCAATGCAACGCCATCTGCGCAAGTACTGCCCGCTCTCGGCGAGT 1982
DB 2968 CTTTCAGCAACCACTTAACGCCATTTGGCGCATTTGATGTTCCACGCTGCTGTCATTT 3027
```

```
QY 1983 CATTGAGCTTGGGACCAATATCAATATTGTTTCACTGTTCAATCCCTCAGACAGGACT 2042
DB 3028 CATTCAATAAGTGTAGGCTGTTGAGGTTGGCTTCGGTAAGCAATTTTCGGCGTCAGACCAC 3087
QY 2043 TACCTGGCTGGC-----GGGGAACCTCTGGAAATTCATCGAGATTTTGT 2090
DB 3088 GAACTTTATTCCTTCACGATGATCCGTAATCTTGGGGAATGCCGTTGAGATTTTC 3147
QY 2091 CGTGCAGCAGCCCTCGCAAGTGTGAGAAAGCAATGACGCCAAGACCAATTTGTCGAG 2150
DB 3148 CGTCAGCAATCTCTGAGCAGAGAGTAAAGGCAATACAGCCACGCCCTTATTTTGA 3207
QY 2151 CTGACTGCAACAAGTTCTCACCGTCATCGCCCGGTTCTTCCACCAAGATTAATGATGG 2210
DB 3208 GGGTATCCAGCA-----GGCGGCTTTTATCCACCCAGCGGTTTCACTAAT 3252
QY 2211 AATAGCTTGGCTGATGATGATCAAGAGCGGCGGCGCTCTCCGCCATGAATCTCAGCGGCT 2270
DB 3253 TGTACGAAGGTTGATGAATTAACAGCGGGAATTTTCCACTCGCGCAGCACTCAACCAIT 3312
QY 2271 CGGCTGAGCTCTGGACCGCTAGGAAGAAATACCCAGTAAAGAGCTTTTCCAGACGCA 2330
DB 3313 TTGCGTTCGCTCTGGCGAGTAAAGAGAGATCCCGACATAAAGCGCTTACCGGTTTGA 3372
QY 2331 CAATGTCACGCAATGCGTACATGTTTCTTCCAAAGAGTATCT 2374
DB 3373 CCGCATGAGCGCAGCGCAGAGCGGTTTCTTCCATCGCGGCTATTT 3416
```

```
RESULT 7
US-08-921-177-20
; Sequence 20, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Heien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TYPE: nucleic acid
; STRANDEDNESS: double
```

us-09-105-117k-1.rni

Thu Mar 18 12:30:59 2004

TITLE OF INVENTION: Probe for Diagnosing Infectious Disease

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/362,577C

FILING DATE: 27-MAR-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien

REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 19036/32420

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

SEQUENCE CHARACTERISTICS:

LENGTH: 5541 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

ORGANISM: Escherichia coli

STRAIN: Clinical Isolate EC-625

US-08-362-577C-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;

Best Local Similarity 50.5%; Pred. No. 2.2e-19;

Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

1623 CATTGTGACGGCCGCTGTCAGCCCAAGGTGTGGCGCTGGATCAAGCTCGTGGCAG 1682

2687 CATTATCCATCTGACGTTTGTGCTCTCTGCAATTTGTGATGGGCTATCTTGACCGACTGAC 2746

1683 TTGTGATGACCGCATTCGCCATCAAACTGATGTTGATGGGTTAGTTTTCGGGGGTTTGG 1742

2747 TCGTCAATATCACTGATCTTTATGCGGGCGGTTCTCGCGCCGTTATTAACAGGTCATT 2806

1743 AATCGGTGGCCCTTCGCCCAAAATTTGATGCGCGCTGGGAAATCTCATCGATCGCCT 1802

2807 TATCGGAAGACCGCTGCCACAGATTCAGCTGCCATATGCTGATCAATCTGGG 2866

1803 CCAACTCGGGGTGAGAAATCTCAAGTGTGATGATGAATCAAGGCTGTGTCCAGCTGTCT 1862

2867 CCAGCTCTCGGTGTGTAATGTCAGATTTATCAGGGCTGCAAGTTCTCTCAAGTTGTC 2926

1863 CAATCTGACGAGCACAATCAATGATCACTGCTCAAGTATCGGCGCTACTCTCTTGTCT 1922

2927 GCGCGGCG-TGGCACTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2967

1923 GCGCGAGCACCACCAATGACGCGCATCTGGCAAGTGAATCTGGCGGCTCTCTGGGCGATGT 1982

2968 CTTTCAGCAACCAAGCTTAACGCCAATTTGGCCCAATTTGTCACCGCTGCTGTCATTT 3027

1983 CATTGAGCTTGGGACCATATCAATATTTTCACTGCTCAATGCTCCCTCAGACAGGACT 2042

3028 CATTCAATAAGTGTAGGCTGTGTGAGTTTGGCTTCGGTAAGCAATTTTCGGCGTCAGACAC 3087

2043 TACCTTGGCTGGC-----GCGGAAACCTCTTGGAAATTCATGATATTGT 2090

3088 GAACCTTTATTCCTTCCATCGATCGATCCGTAATCTTGGGAAATCGCGTTGAGATATTTTC 3147

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE:

ORGANISM: Escherichia coli

STRAIN: Clinical Isolate EC-625

US-08-921-177-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;

Best Local Similarity 50.5%; Pred. No. 2.2e-19;

Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

1623 CATTGTGACGGCCGCTGTCAGCCCAAGGTGTGGCGCTGGATCAAGCTCGTGGCAG 1682

2687 CATTATCCATCTGACGTTTGTGCTCTCTGCAATTTGTGATGGGCTATCTTGACCGACTGAC 2746

1683 TTGTGATGACCGCATTCGCCATCAAACTGATGTTGATGGGTTAGTTTTCGGGGGTTTGG 1742

2747 TCGTCAATATCACTGATCTTATGCGGGCGCGGTTCTCGCGCCCGTTTATTAACAGGTCATT 2806

1743 AATCGGTGGCTTCCGCCAATGTTGATGCCGGCTGTGGGAAATCTCATCGATCGCCT 1802

2807 TATCGGAAGACCGCTGCCACAGATTCAGCTGCCATCGCGGATATGCTGATCAATCTGGC 2866

1803 CCAACTCGGGTTCAGAAATCTCAAGTGTGTGATGATCAAGCTGTGTTCAGCTGCT 1862

2867 CCAGCTCTCGGTGTAATGTCAGATTTATTCAGCGCTGCGAGTTCTCTCAAGTTGTC 2926

1863 CCACTGACGAGCACAATCAATGATGTTGATGATGATCAAGCTGTGTTCAGCTGCT 1922

2927 GCGCGGC-TGGCACCACCAATACCGAAGTCACGCGATCA-----T 2967

1923 GCGCGAGCACCATGACGCGCATCTGCGCAAGTGAATGCTCGCGCTTCTCGGGGATGT 1982

2968 CTTTCAGCAACCAAGCTTAACGCCAATTTGCGCCATTTGATGTCACGCTGCTGTCGCAATT 3027

1983 CATTGAGCTTGGGACCATATCAATATGTTGATGATGATCAATGCTCAATGCTTCAAGGACT 2042

3028 CATTCAATAAGTGTAGGCTGTTGAGGTTGGCTTTCGGTAAGCAATTTTCGGCGTCAGACAC 3087

2043 TACCTTGGCTGGC-----GCGGAAACCTCTTGGAAATTCATGATATTGTT 2090

3088 GAACCTTTATTCCTTTCAGATGATTCGCGTAATCTTGGGAAATGCGGTTGATATTTC 3147

2091 CCGTGACGAGCCCTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCAATTTTGGCAG 2150

3148 CGGTGACGATCTCTGACGACGAGGATGAAGCAATACAGCCCAAGCTTATTTGCA 3207

2151 CTGACTGCAACAAGTTCTTACCGTCAATCGCCCGGTTCTTCCACCCCAAGATTAATGATGG 2210

3208 GGGTATCCAGCA-----GCGCGCTTTTATCCACCGCGGTCAGTAAT 3252

2211 AATAGCTTGGCTGATGATCAGAGCGGCGAGCCCTCTCGCGCATGAATCTAGCGCGCT 2270

3253 TGTACGAAGTTGATGAATTAACAGCGGAATTTTCACTGCGGAGCACTCAACCAATT 3312

2271 CCGCTGTGAGCTCTGGAACCGTAGGAAGAATAACCGATGAAGAGCCCTTCCAGAGCGAA 2330

3313 TTTCGCTCGCTCTCGCGAGTAGAGAGATCCCGACATAAAGCGCTTACCGTTTGTGA 3372

2331 CAATGTACGCAATCGGTACATGTTTCTTCCAAAGGATATCT 2374

3373 CGCATGAGCAGCGCAGAGCGGTTTCTTCCATCGCGGTAATT 3416

RESULT 8

US-08-362-577C-20

; Sequence 20, Application US/08362577C

; Patent No. 5807673

; GENERAL INFORMATION:

; APPLICANT: Onno, Teuneya

; APPLICANT: Matsuhisa, Akio

; APPLICANT: Uehara, Hirotsugu

; APPLICANT: Eda, Soji

RESULT 9
 US-08-920-828-20
 ; Sequence 20, Application US/08920828
 ; Patent No. 5853998
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohno, Tsuneya
 ; APPLICANT: Matsuhisa, Akio
 ; APPLICANT: Uehara, Hirotsugu
 ; APPLICANT: Eda, Soji
 ; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/920,828
 ; FILING DATE: 29-AUG-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/362,577
 ; FILING DATE: 27-MAR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures, Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 19036/32420
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5541 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Genomic DNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Escherichia coli
 ; STRAIN: Clinical Isolate EC-625
 US-08-920-828-20

Thu Mar 18 12:30:59 2004

ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5370
Query Match
Best Local Similarity 51.5%; Pred. No. 5.2e-14;
Matches 252; Conservative 0; Mismatches 216; Indels 21; Gaps 2;
QY 1889 CTGTCACGGTATCGCGCGTACTCTCTCTGCTCGCGCAGCACCACCATGCAACGCGCATC 1948
DB 998 CTGGCCGCAATCAGACCGAGGTGACTTCTCTCGCGCAACACCCAGCGCGCATC 939
QY 1949 TGGCGAAGTGAATCGCGCGTCTCTGGGCGATGATGAGCTTCGGGACCATCAATA 2008
DB 938 TGGCAGAGCTTCTGCGCCACGCGTTCGCGCTGCTGTTAAGCTGACGAAATTTTCTAAT 879
QY 2009 TTGTTACAGTTCAACATGCCCTCAGACAGGACTTACCTGGCT---GGCGCGGGAACCC 2065
DB 878 CGGCTGGAGTCAGTTGCTCGGCTCGAGAAACGATGCTGCTGGCGCGCGGAGTCG 819
QY 2066 TCTGGAATTCATCGAGATATTTGCTCGTGAAGCGCCCTGCGCAAGTGGTGAAGAACA 2125
DB 818 GCGGATGCGCTGAGATAGCGGTGGTAAAGTTCGCGCGCGCGGAGCGTGAAGACGCT 759
QY 2126 ATGACGCGAAGACCATTTGGCAGCTGACTGACACAGTTCCTCACCGTCATCGCGGT 2185
DB 758 ATGCTTCAATCCCTTCGCTGCGAGAAATCCAGCAATCC-----T 717
QY 2186 TCTTCCACCCACGATTAATGATGAATAGCTTGGCTGATGAATCAGAAGCGGCGAGCCC 2245
DB 716 TCTTCCACCGCGGTTCGAATCATGATGAGCGAGGCTGGTGAATGATGACGCGGTGCGG 657
QY 2246 TCTTCCGCAATGACTAGCGCGCTCGCTGTGAGCTCTGACCGGTGAGAGAAATACC 2305
DB 656 AGGTCGTTAAGGATCTTCAACCGCTCCGCGCTCGCGCGCTGGCGAGAGATGTTGAGATCCG 597
QY 2306 AGCTAAGAGCTTCCAGACGCAAGTGTACGATGCTGATGCTGATGCTTCTTCCAAA 2365
DB 596 ACATATAGCGCTTACCTCGAGAACAGATGATGACGCGACGATCGTCTTCCATCAGC 537
QY 2366 GGAGTATCT 2374
DB 536 GCGCTCTGT 528
RESULT 11
US-09-489-039A-4735/c
; Sequence 4735, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4735
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4735
Query Match 3.4%; Score 80.2; DB 4; Length 915;
Best Local Similarity 44.9%; Pred. No. 1.1e-13;
Matches 393; Conservative 0; Mismatches 473; Indels 9; Gaps 2;
QY 87 GCGCAATCCCTGATGCTGTCATCAACGAGCGGTCTGTGAGCTCTAGCTAGATCTAG 146
DB 907 GACGACGACCTTGTGTCGCTAGTCAATGAGCGCATCGGTACCGCGCGCATCATGCGGC 848
QY 147 ATTCAGGCGCCATGTTGCCAATA---CATCGGTGTCAATGGGTATCTCATCGAGGA 203

DB 847 TTTCGGGCGAGCGTGCAGTAGAGCATGCGGCGCTGGAACAGGCCCGGGGTGAGAT 788
QY 204 GGATCACTCTCTCTGCTTTTAGCATGGGAGCAGCTTGGGTTTCGGGAAGAAGTCCCCAAC 263
DB 787 CGATAAGTTTCGCGCGTGTGAGCTCTTTTTCGATCTGAGATGCGGGATCATACAGCAGG 728
QY 264 CAAGGCTCGGCGAATTCCTTCAACAAACCTTCGCGCGAGCGGCAATGATGATACGCGCC 323
DB 727 TGGTCCCTGGCGCGCCAGCTGACGAAGCTTCGACGAGTTCGACGATGTGGCAGGGGA 668
QY 324 TGGCCCCCAACAGGACCATTCGAGCGCGCCCTTCAGGTCAAGTCAAGTCTTGAAGACATCTTTGG 383
DB 667 GCTTCCCGCGCGCAGGTTCGAAGTTTGTGTCAGGAAACGCTTGATGCATATCTGTCAGAT 608
QY 384 GACGAGCGTGAAGACGGC-----ATCGCAGGCCAATCTAGTTTTCCTCATCAACCATGT 437
DB 607 GGTGAAAGGCGACGACCGCGGCTTAAAGCAGCGCGCAACGCTCAGCGCAITCGGGAAT 548
QY 438 AGGCATCCCGCAATGAGGGGTTGCAATGGCCAAAGTGGCGCATGGTTCCAGTTCTACTA 497
DB 547 AGCGTTCGCGCAACTCTTTTATGCGCAAAACAGGTAAATCGAGCGCGCCCGAGCTGATCCA 488
QY 498 CTTCAATTCGCGCAGCGGATAGCTTCACGGGTTACCGTCTCTAAACATCTCCAGGCC 557
DB 487 CCAGGCGAGCTTGGCAGCGCTTGGGCTGAGATCTACTGCGGCCACTACTTTCGCGCAGGC 428
QY 558 GCAGCAAGGATATGTTGGCGCTTCACTTTCGAAGCGAGCGGTGAGCGTGTCTCCACCCC 617
DB 427 GAGCGCGCTCTGGGTGCGGTTTCACTTTCACCTGCGAGGTTGAGACGAAATAGGGGAGT 368
QY 618 AAGAAGCTACTCTGTTGAACACGCGGAGGAAACCATGTGATAGCGATCTGCGTTGATGG 677
DB 367 CTGACAGAACTGTTGGCGCGCGCGCAGCAGCGCTGCGCAGACTGTCGCGGTTTCCCG 308
QY 678 CGATGTTAAGCGGATTCAGCAAGCGGCTCCAGATAGTTGCGCTTTAGTTTCTGCTGCA 737
DB 307 CCAGCGCAGCAGCAACGCGGTAGAGCGGTTGTTTCATCGCCCGCAGCACTCTCTCTCCA 248
QY 738 GCAACACCATTTTCCCGCTGCTTGCACAAAGGACTTCACCGCGCTTCGTTGCTTGGCGG 797
DB 247 GAGTTTCAACTGCGCGCAACAGCAGCGAGTTCCTGCTTGTCTGCGTGGCGCGCGG 188
QY 798 GTTGGTTCGCGATACCAACACTGACCGCAGTGTCTCGAGAGCTTTAAACCGCTGAC 857
DB 187 GCAAGCTACGACCAACGCGCTGCGCAACATATTTTCCAGCTGCTTAATACGCTGTG 128
QY 858 TCACGCGCGGCGGAAATGGAAGCGCTAAGGAGCGGCTTCGAAGCTGCTCTCAATCA 917
DB 127 AGACGCGGACTGCGGTAATACATAGCTTTCGCGCGCGGCTCAAAATCCTCGTTCCCTAA 68
QY 918 TGATTGAGGCAAAAGTGTCCAGTTGAATGGGTTTC 952
DB 67 TCACCGCATCAGCGCTTGTAGTGTCTCTGTAGTCC 33
RESULT 12
US-09-543-681A-791/c
; Sequence 791, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAB
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 791
; LENGTH: 936
; TYPE: DNA

ORGANISM: Proteus mirabilis
US-09-543-681A-791

Query Match 3.1%; Score 73.4; DB 4; Length 936;
Best Local Similarity 44.9%; Pred. No. 1.2e-11;
Matches 324; Conservative 0; Mismatches 391; Indels 6; Gaps 1;
QY 206 ATCACTCTCTCTCTTTTATGATGGGAGCAGCTTGGGTTTCGGGAGAGATCCCAACA 265
Db 800 ATCAATTCACCACTTTTATGCTCATCGGCAATTTGTAGATGAGGAGCATACAACTGT 741
QY 266 AGCCCTCGCGGATTTGCTTACCAAACTTTCGGCGAGCGGACAAATGATAGCGCCCTG 325
Db 740 GAGCTTGTGTTAGCTTAATCAAAAGGCTTCGGAAGATTAACAAATATGCGAGGCGACA 681
QY 326 CGCCCCACAGGACCATCGAGCGCCCGCTCCAGGTCACGGTCTTTGAAGACATCTTTGGGA 385
Db 680 CTACCGGAGATAAACCGAGTTTTCCTGTAAAGAGCCCTGATGCTATCATCAAGATGG 621
QY 386 CGAAGCGTGAAGACGGG-----CATGCGACCCCAATCTAGTTTCCCATCAACCATGTAG 439
Db 620 TCAATGCGCACAGCGGGGGCTTTTAAACAGAGATGATTTAGTGACCCCAATTTGCAAAATAG 561
QY 440 GCATCCCGCAATGAGGGGGTTGCAATGCGCAAGTGGCGCATGTTCCAAAGTTCTACTACT 499
Db 560 CGTTGGCGAAATACAGGGGATGCCCAAAAGATAGTCTTAATGCGCCTTAATGTCGACT 501
QY 500 TCATATCCCGCACCGGATAGCTTCACGGTTTACCGCTCTTAAACATCTTCAACCGCGC 559
Db 500 AGACAGTAGGCAATGCTTGGGGCTGAATCTGATAGACCAACCACTTCAACAGCTCTT 441
QY 560 AGCAAGGATAATGTTGGCGTTTCACTTCCAGCGAGCGGTGAGGTTGCTTCAACCCCA 619
Db 440 AATCGCTCTCTGCTACGAGTTTCACTCTTCTAATTAAGCCGAGTAGGAAGTTGT 381
QY 620 GAAGCTACCTCTGCAACAGCGGAGAACCATGTGATAGCGAATCTGCGTTGATGGGG 679
Db 380 GTTAATACCGGTGCAAGAGGCAATTAACAGTAGTAACTATTCGCGTTGACAGCT 321
QY 680 ATGTTTAAAGGATTTTACGAGGCGTCCAGATAGTTGCGCTTTAGTTTCTGCTTGCAGC 739
Db 320 AAAGAGGTAAAGAGGCGTCGAGCCACTATTTTTCATCACTTAACCACTGCTCTTCTAAC 261
QY 740 AACACCATTTTCGCGCTGCTTGGCAAGGACTTCAACCGCTTGGGTTGCTTGGCGGT 799
Db 260 ATTTGCACTGTAGTCAATGAGTAGTCTTTGCGCCCTGCTCTGCGGTTGGGCGGA 201
QY 800 TGGGTGCGGATACCAACATCTGACCCACGCTGATGCTCGAGAGCTTTAAACGCGTGACTC 859
Db 200 ACCGTAGTACTAATAAAGGTTGTCAAATAGTCTCAAGTTGTTTATTTCTTTGTGA 141
QY 860 ACCCGCGAGGGGAAATGAAAGGCTTAAGGAGCGGCTTCAAGCTGCTTCAATCATG 919
Db 140 ACGGAGATTGCTTAATACATAGTTTGTGGCGGCTGTTCAAAACCAACGCTCTCGAATA 81
QY 920 A 920
Db 80 A 80

RESULT 13
US-09-252-991A-7644/c
Sequence 7644, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7644
LENGTH: 900
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7644
Query Match 2.7%; Score 65.2; DB 4; Length 900;
Best Local Similarity 46.3%; Pred. No. 3.2e-09;
Matches 412; Conservative 0; Mismatches 453; Indels 25; Gaps 5;
QY 87 GCAGCAATCCCTCGGATTTGCTCATCAACGAGCGGCTCTGTGAGTCTTAGCTAGAGATCTAG 146
Db 886 GCACCAAGCGTTCGCCGGCGGCGGAGCAGTCTCGTGTAGCGACGAGCAGTTCGCG 827
QY 147 ATTCCAGGCGCATCGTTGCGCAATATCATCGTGTGTCTATGGGTATCTCATCGAGAG-- 204
Db 826 CGCGTTTGGCGAGTAGTGGCAGTACACGCTAGCTCGATGACCTGGCGGCGAGAGCT 767
QY 205 -GATCATTCTCTCTTTTATGATGGAGCAGCTTGGGTTTCGGGAGAGAGTCCCCAAC 263
Db 766 CGACCACTTCCAGCGCGCCAGCTTCGCTTTCGCTTCCGCTCCGCGACCGAGCCCGC 707
QY 264 CAAGGCTCGCGCAATTCGCTTCCAAACCTTCGCGCGAGCGGACCAATGGATAGCGGC 323
Db 706 GAGACCGCGCGGCTGAGCGGAGCAACCTTCGAGGAGCGGCGAGAGTGGTGGATGA 647
QY 324 TGCGCCCAACAGACCAATCGAGCGGCGGCTTCAGGTTCAGGTCTTGAAGCAGCATCTTTGG 383
Db 646 AACCGCTCTCGA-----CGCGAGGTCTTGAAGAGCGGTGTTCGACAGGTTCGTCG 593
QY 384 GACCGAGCTTAAGACGCGGATTCGAGCGCAATCTAGTTTCCATCAACCAATGATAGCAT 443
Db 592 GCGCGAAACAGATTCGCGCGGACGCGCGGCGAGCGCGCGCTCGA.CCCCGCGGGAAGT 533
QY 444 CCGCG-----AATGAGGGGGTTCGAATGCGCAAGTGGCGCATGGTTCAGATTTCTACTA 497
Db 532 GCGCGCAATGAATTCGCGGCTGCGCGAGCGCGCATACGATGGCTTCGAGCAGCAGCG 473
QY 498 CTTCAATCCCGCGAGCGGATTTAGCTTCAAGGTTTACCGCTTCTTAAACATCTCCACGCC 557
Db 472 AACCGCGCGCGCGACCGCGCGCGCTACCGCAGACGAGCGCGGCTTCGCGCGCGC 413
QY 558 GCAGCAGGATATCTGTGCGCTTCAATCTTCAAGCGAGCGTTCGAGGTTCCTCCACCCC 617
Db 412 GCATGCGCTTGAAGCGGCTTCTGCTTCTTCCACAGAGTTCAGGAGCAGCAGCGCGCT 353
QY 618 AAGAGCTTACTCTGTTGAACACGCGGAGGAAACCATGTGTGATAGGAAATCTGCTTGA--TG 676
Db 352 CGCGCGCAATTCGCCACCGCGCGCGCCACCGAGTCCGAGGCTATCGCGGTTGAGCA 293
QY 677 GCGATGTTTAAAGGATTTACGAGGCTTCAGAGCGCTTCAGATGTTGCG-----CTTAGTT 727
Db 292 GCGATGCGCGCGCTTCGCGGCGCCCGCTTCATCCAGGTTTCGCGACCCAGCGCTGAGG 233
QY 728 TCTGCTTCGAGCAACACCATTTTCGCGCTGCTTTCGCAAGGACTTTCACCGCTTCGCTT 787
Db 232 TCGCCCTTCAGCAGCGCGCACCTTCGCTGTCGATGTCGATGTCGAGAGCTTTA 173
QY 788 GCTTTGGCGCGTTCGCGGATACCAACATCTCGACCCAGCTGATGCTTCGAGAGCTTTA 847
Db 172 GGTTCGCGCGCTCTTCGCGCACACCGAGCGGCTGCGCGGAGCGCTTCGAGAGCTTTG 113
QY 848 AGCGCTGATCTACCGCGAGGAGGAAATGGAAGGCTTAAGGAGCGCTTCGAGAGCTG 907
Db 112 ATCCGCTCGGATACCGCGGATTTGAGACACCCCGCGCGCTGTGCGGCGCGCTCGAAACCG 53
QY 908 CTTTCATCAATGATTGAGAGCAAGTGTCCAGTTGATGGGCTTCATGA 957
Db 52 CCGTCTCCACCTGCGCGCAGCGCGCGGCACTTGTAGTCGAACAA 3

Thu Mar 18 12:30:59 2004

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7908
LENGTH: 699
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7908

Query Match 2.6%; Score 61.8; DB 4; Length 699;

Best Local Similarity 54.6%; Pred. No. 2.9e-08;
Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;

QY 1455 TGATGGCAATCGTGTGACCTGGTTGAACCCGAATCGCTATTGGACGGGTTGTTTGA 1514
DB 434 TGGCGGCACTGGCGGTCACCCCTGCTCAACCCCGACGCTCTACTCGATACCGTACTCTGA 493
QY 1515 TCGGCGGCGTGGCGGCAATACGGCGACACCGGACGGTGGATTTTCGCGCTGGCGGT 1574
DB 494 TGGTTTCGCTCGGCGCCGACGAGG-----CCGCGCGGCGGCGCTATGCCCCCTCGCGCGG 547
QY 1575 TCGCGGCAAGCTGATCTGTTCCGCTGGTGGTTTCGCGCGACGACGATTTGTCACGCC 1634
DB 548 CCAGCGCCTCGCTGATGTGGTTCTTCGCCCCCTCGCGCGCATGGCTGGCCCCCT 607
QY 1635 CGCTGTCCAGCCCCCAAGGTGTGGCGCTGGATCAACGTCGTGGCGAGTTGTGATGACCG 1694
DB 608 GGTGGCGGCGCGCGCCACCTGGCGCTGCTCGACCTGATGGTGGCGGCATGATGCTGG 667
QY 1695 CATTTGCCATCAAACTGATGTTGATGGTTTA 1725
DB 668 GCATGGCGCGCAACTGCTGTTCCGCGGATA 698

Search completed: March 16, 2004, 03:00:33
Job time : 198.837 secs

RESULT 14

US-09-724-623-24/c
Sequence 24, Application US/09724623
Patent No. 6476209
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Lubbers, Mark W
APPLICANT: Dekker, James
TITLE OF INVENTION: Polynucleotides, materials incorporating
TITLE OF INVENTION: them, and methods for using them.
FILE REFERENCE: 1048U1
CURRENT APPLICATION NUMBER: US/09/724,623
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 124
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 1002
TYPE: DNA
ORGANISM: Lactobacillus rhamnosus
US-09-724-623-24

Query Match 2.7%; Score 64; DB 4; Length 1002;

Best Local Similarity 47.9%; Pred. No. 7.8e-09;
Matches 279; Conservative 0; Mismatches 270; Indels 33; Gaps 2;

QY 1785 AATCTCATCGATCGCTCCCACTCGCGCTCAGAAACTCCCAAGTTGTTGAGTGAATCAA 1844
DB 994 AATATCATCAATTTGTTGATCTCTTCAGCAGTAAGGTCAAGTGTTCGGTTGCTTAA 935
QY 1845 GCCTGTGTCAGCTCTCACTGACGAAGCACCATCAATGATGCTGTCAGGTATCCG 1904
DB 934 GGTATCTGAAGGTGTTCACTGAGTGTGTCCTCCATGATCACATTTGTGAAC----- 881
QY 1905 CGCGGTACTCTCTGCTGCGCAGCACCCTGACAGGCCATCTGCGAAGTACTGCC 1964
DB 880 -----CGGATCCCGTAACACACCGCCAGGCCATTGACTCAGGTTTGT 833
QY 1965 CGGTTCTGCGGATGATGATGAGCTTGGGACCATATCAATATTTTCAGGTTCAACA 2024
DB 832 CAGATCATGCGAATTTCTTAAGCGATTAGTTGCTTAACACAGCCTCTTTGCCCT 773
QY 2025 TGCCTCTCAGACAGGACTTACCTGCTGGCGGCGAACCCTCTGGATTCATCGAGAT 2084
DB 772 TAGCAAAAGTGGCTTGTGGTGGATGGATTTTGAAGTATCGGAATTCCTTTAGT 713
QY 2085 ATTGTGCTGAGCAGCGCTGCGAAGTGTGAGAAAGCAATGACGCCAAGACCATTTGT 2144
DB 712 AGCGATCTGATAACAGCCTTCTGATACGGTCCGTATGCATCAACCGCGCACCATCAG 653
QY 2145 TGGCAGCTGACTGCAACAAAGTTCTCACCGTATCGCCGGTTCCTCCACCCACGATTAA 2204
DB 652 CTTTAAATGATCGATCAA-----GCCGACGTTTTCAGCGGTGCGATTAA 608
QY 2205 TGATGGAATAGCTTGGCTGATGATCAGAAGCGGCGACCTCTCTCCGATGATGATCAG 2264
DB 607 ACATGTTGTAACCTGATTGATTCAGTACAAAGCGGTGTCAGATCTTTAAAGATCGCA 548
QY 2265 CGGCTCGCTGAGCTCTGAGCGGTAGGAGAAATACACGCTAAAGAGCCTTTCCAG 2324
DB 547 TTGCTTCTTTGCTGCTGCTATCATAGTTGGAAATACCAATATACACGCTTACCGT 488
QY 2325 AGCAACAAATGTCAGCAATGCGTACATGTTTCTTCCAAAG 2366
DB 487 CAGCAGCGTTTGATCCAGCGCATTAACCGTCTCTTCCAAAG 446

RESULT 15

US-09-252-991A-7908
Sequence 7908, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	2374	100.0	2374	14	US-10-196-232-24		Sequence 24, Appl
C	2374	100.0	330940	9	US-09-738-626-1		Sequence 1, Appl
	2374	100.0	330940	9	US-09-738-626-1		Sequence 47, Appl
3	993	41.8	993	10	US-09-746-660A-47		Sequence 3456, Ap
C	870	36.6	870	9	US-09-738-626-3456		Sequence 51, Appl
	822	34.6	822	10	US-09-746-660A-51		Sequence 7, Appl
5	822	34.6	822	10	US-09-746-660A-51		Sequence 3455, Ap
6	711	29.9	711	14	US-10-166-142-9		Sequence 9, Appl
7	708	29.8	708	9	US-09-738-626-3455		Sequence 3454, Ap
C	627	26.4	627	9	US-09-738-626-3454		Sequence 659, App
	627	26.4	627	9	US-09-738-626-3454		Sequence 17521, A
10	485.2	20.4	993	12	US-10-627-476-669		Sequence 17521, A
C	305.8	12.9	1095	12	US-10-282-122A-17521		Sequence 35374, A
C	302.8	12.8	879	12	US-10-282-122A-17523		Sequence 38826, A
C	164.4	6.9	1026	15	US-10-369-493-35374		Sequence 38826, A
C	164.4	6.9	1026	15	US-10-369-493-38826		Sequence 38124, A
C	164.4	6.9	1028	15	US-10-369-493-38124		Sequence 38124, A

121	GTCTGTGAGTCTAGTAGAGATCTAGATTCACGCGCCATCGTTGCCAATACATCGGTGT	180	TTTGTCCAAATCGCGCGCATGCTGCTCGATATTTATGCTGCTGGGTGGCATCGTTACCT	1260
121	GTCTGTGAGTCTAGTAGAGATCTAGATTCACGCGCCATCGTTGCCAATACATCGGTGT	180	TTTGTCCAAATCGCGCGCATGCTGCTCGATATTTATGCTGCTGGGTGGCATCGTTACCT	1260
181	GTCAATGGGTATCTCATCGAGAGATCATCTTCTCTGCTTTTAGCATGGAGCAGCTTG	240	GTATGTTGGCGGTCTATGGCAGCAAGACCGCATGACAAACAAAGGTGGAAGCGCACCA	1320
181	GTCAATGGGTATCTCATCGAGAGATCATCTTCTCTGCTTTTAGCATGGAGCAGCTTG	240	GTATGTTGGCGGTCTATGGCAGCAAGACCGCATGACAAACAAAGGTGGAAGCGCACCA	1320
241	GGTTTCGGAGAGTCCCAACCAAGGCTTCGCGCAATTTGCTTCCAAAACTTCCGC	300	GATCATTTGAGAAACAGAACCAACCGTCCGATGACACGCTTTTGGCGGTTTCGGCGGT	1380
241	GGTTTCGGAGAGTCCCAACCAAGGCTTCGCGCAATTTGCTTCCAAAACTTCCGC	300	GATCATTTGAGAAACAGAACCAACCGTCCGATGACACGCTTTTGGCGGTTTCGGCGGT	1380
301	CGACGGACAATGATAGCGGCTCGCGCCCAACAGGACCATCGACGCGCTTCAGGTTC	360	GGCCACTGACACGCGCAACCGGTGCGGTGAGGTGAGGTGATAGCAGCGGTTCG	1440
301	CGACGGACAATGATAGCGGCTCGCGCCCAACAGGACCATCGACGCGCTTCAGGTTC	360	GGCCACTGACACGCGCAACCGGTGCGGTGAGGTGAGGTGATAGCAGCGGTTCG	1440
361	ACGGTCTTGAAGCAATCTTTGGGACCGAAGGTAAAGACGGCATCGACCCCAATCTAG	420	GTTAAAGCCCATTTGATGAGCAATCTGCTGACCTGCTGAAACCCGATGCTATTGGA	1500
361	ACGGTCTTGAAGCAATCTTTGGGACCGAAGGTAAAGACGGCATCGACCCCAATCTAG	420	GTTAAAGCCCATTTGATGAGCAATCTGCTGACCTGCTGAAACCCGATGCTATTGGA	1500
421	TTTCCCATCAACCATGAGCATCCGCAATGAGGGGTTGCAATGGCCAAAGTGGCGAT	480	CGCGTCTTGTATTCGGCGCGCTCGCGCGCAATGAGCGGACACCGGACCGGTGATTT	1560
421	TTTCCCATCAACCATGAGCATCCGCAATGAGGGGTTGCAATGGCCAAAGTGGCGAT	480	CGCGTCTTGTATTCGGCGCGCTCGCGCGCAATGAGCGGACACCGGACCGGTGATTT	1560
481	GGTTCCAAAGTCTTACTACTTCAATCCCGCCACGCGATTAGCTTACCGGTTC	540	CGCGCTGCGCGGTTTCGGCGCAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
481	GGTTCCAAAGTCTTACTACTTCAATCCCGCCACGCGATTAGCTTACCGGTTC	540	CGCGCTGCGCGGTTTCGGCGCAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
541	TAAACATCTCACGCGCAGCAGGATATGCTGGCTTCACTTCCCAAGCGCAGCT	600	AGCATTTGACGCGCTTCGCGCGCTCCAGCGCTGCGCTGCTGCTGCTGCTGCTGCTG	1680
541	TAAACATCTCACGCGCAGCAGGATATGCTGGCTTCACTTCCCAAGCGCAGCT	600	AGCATTTGACGCGCTTCGCGCGCTCCAGCGCTGCGCTGCTGCTGCTGCTGCTGCTG	1680
601	GAGGTTGCTCCACCCCAAGAGCTACTCTGTTGAACACGGGAGAAACCATGAGTAG	660	AGTTGCTGATCAGCGCATTTGGCCATCAAACTGATGTTGATGGGTAGTTTCGCGGT	1740
601	GAGGTTGCTCCACCCCAAGAGCTACTCTGTTGAACACGGGAGAAACCATGAGTAG	660	AGTTGCTGATCAGCGCATTTGGCCATCAAACTGATGTTGATGGGTAGTTTCGCGGT	1740
661	CGAATCTCGCTGATGCGGATGTTAAAGGATTTACGAGGCTTACGAGGTTCGCG	720	GGATTCGGTGGCTTCGCGCAAAATGTTGATGCGCGCTCGTGGGAATCTCATCGATCG	1800
661	CGAATCTCGCTGATGCGGATGTTAAAGGATTTACGAGGCTTACGAGGTTCGCG	720	GGATTCGGTGGCTTCGCGCAAAATGTTGATGCGCGCTCGTGGGAATCTCATCGATCG	1800
721	TTTAGTTCTGCTGACGACACCATTTTCGCGCTGTTGACACAGGACTTCAACCGC	780	CTCCAACTCGCGCTCAGAAATCTCCAAAGTTTGTAGTGAATCAAGGCTGTTGTCAGCTG	1860
721	TTTAGTTCTGCTGACGACACCATTTTCGCGCTGTTGACACAGGACTTCAACCGC	780	CTCCAACTCGCGCTCAGAAATCTCCAAAGTTTGTAGTGAATCAAGGCTGTTGTCAGCTG	1860
781	TTTCGTTGCTTTGGCGGTTGGTGCGGATACCAACTCGACCCACGATGATCGAG	840	CTCAACTGACGAGCACCACATCAATGACCTGCTCAGGTATCCGCGCTGCTCTCTCTG	1920
781	TTTCGTTGCTTTGGCGGTTGGTGCGGATACCAACTCGACCCACGATGATCGAG	840	CTCAACTGACGAGCACCACATCAATGACCTGCTCAGGTATCCGCGCTGCTCTCTCTG	1920
841	AGCTTTAACGCGCTGACTCACCGCGAGGGGAAATGGAAGGCTAAGGAGCGCCTTC	900	CTCGCGAGCACCACATGCAAGCGCATCTGCGCAAGTACTGCGCGCTTCCTGGCGAT	1980
841	AGCTTTAACGCGCTGACTCACCGCGAGGGGAAATGGAAGGCTAAGGAGCGCCTTC	900	CTCGCGAGCACCACATGCAAGCGCATCTGCGCAAGTACTGCGCGCTTCCTGGCGAT	1980
901	GRAGTGGCTTCAATCATGATGAGCAAAAGTGTCCAGTTGAATGGGGTTCAAGAGCT	960	GTCATTGAGCTTGGGACCATATCAATATTTGTTACGTTTCAACATGCCCTCAGACAGGA	2040
901	GRAGTGGCTTCAATCATGATGAGCAAAAGTGTCCAGTTGAATGGGGTTCAAGAGCT	960	GTCATTGAGCTTGGGACCATATCAATATTTGTTACGTTTCAACATGCCCTCAGACAGGA	2040
961	ATATTAACCAATGTTAAGAACCAATCAATTTTACTTAAGTATCTCCATAGTTCAGATGT	1020	CTTACCTGGCTGCGCGGAAACCTCTGGAATTTCCATCGAGATATTTGTCGGTAGAG	2100
961	ATATTAACCAATGTTAAGAACCAATCAATTTTACTTAAGTATCTCCATAGTTCAGATGT	1020	CTTACCTGGCTGCGCGGAAACCTCTGGAATTTCCATCGAGATATTTGTCGGTAGAG	2100
1021	GATCATGGAAATCTTATACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCATCGG	1080	GCCTTCGCAAGTGGTGGAGAAACATGACCCCAAGACCATTTGTTGGCAGCTGACTGCAA	2160
1021	GATCATGGAAATCTTATACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCATCGG	1080	GCCTTCGCAAGTGGTGGAGAAACATGACCCCAAGACCATTTGTTGGCAGCTGACTGCAA	2160
1081	ACCGAGAAATGATCTGATGATTAACAGGAAATTAAGCGGAGGACTTCAATCGGTTCT	1140	CAAGTTCTCACCGTCACTCGCGCTTCCTCCACCAACGATTAATGATGGAATAGCTTGG	2220
1081	ACCGAGAAATGATCTGATGATTAACAGGAAATTAAGCGGAGGACTTCAATCGGTTCT	1140	CAAGTTCTCACCGTCACTCGCGCTTCCTCCACCAACGATTAATGATGGAATAGCTTGG	2220
1141	TCTGCTGTTTAAATTTCTGAGCTCTTTTGTGATCGCGGACCTTTGGCGGTGATCT	1200	CTGATGAATCAGAGCGGCGGCTTCCTCCGCCATGAATCTCAGCGGCTCCGCTGTGAG	2280
1141	TCTGCTGTTTAAATTTCTGAGCTCTTTTGTGATCGCGGACCTTTGGCGGTGATCT	1200	CTGATGAATCAGAGCGGCGGCTTCCTCCGCCATGAATCTCAGCGGCTCCGCTGTGAG	2280
			CTCTGGACCGTAGGAGAAATACCCACGTAAAGAGCCTTTTCCAGACGCAACATGTCAG	2340

Db 2281 CTCTGACCGTAGGAAGAAATACCAAGTAAGAGCGCTTTCCAGACGCAAAATGTCAGC 2340
QY 2341 CAATGGTACATGTTCTTCCAAAGGAGTATCT 2374
Db 2341 CAATGGTACATGTTCTTCCAAAGGAGTATCT 2374
RESULT 2
US-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US2002019705A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Query Match 100.0%; Score 2374; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATTTCTGAAGGTGTTACTCTGCTGGCCCAATTCCTCGGGCGAAGAGTGAAGAAC 60
Db 1329968 CCAATTCGTGAAGGTGTTACTCTGCTGGCCCAATTCCTCGGGCGAAGAGTGAAGAAC 1329909
QY 61 CCGAACCTTTTCAGAGTAACTAAGCCGCAATCCCTCGATTGCTGCATCAAGAGCGGC 120
Db 1329908 CCGAACCTTTTCAGAGTAACTAAGCCGCAATCCCTCGATTGCTGCATCAAGAGCGGC 1329849
QY 121 GTCTGTGAGTCTAGCTAGAGATCTAGATTCCAGGGCCCATCGTTGCCAATACATCGGTG 180
Db 1329848 GTCTGTGAGTCTAGCTAGAGATCTAGATTCCAGGGCCCATCGTTGCCAATACATCGGTG 1329789
QY 181 GTCAATGGGTATCTCATCGAGAGATCACTTCTCTCTCTTTAGCATGGGAGCAGCTTG 240
Db 1329788 GTCAATGGGTATCTCATCGAGAGATCACTTCTCTCTCTTTAGCATGGGAGCAGCTTG 1329729
QY 241 GGTTCGGGAAGAGTCCCAACCAAGGCTCGGGGAAATGCTCTCACCRAAACCTTCGC 300
Db 1329728 GGTTCGGGAAGAGTCCCAACCAAGGCTCGGGGAAATGCTCTCACCRAAACCTTCGC 1329669
QY 301 CGACGGGACAAATGATACCGGCTCGCGCCCAACAGGACCATCGACGCGCCCGTCCAGGTC 360
Db 1329668 CGACGGGACAAATGATACCGGCTCGCGCCCAACAGGACCATCGACGCGCCCGTCCAGGTC 1329609
QY 361 ACGGCTTTGAAGCACATCTTTGGGACCGAAGCGGTAAAGAGGGGATCGAGGCCAATCTAG 420
Db 1329608 ACGGCTTTGAAGCACATCTTTGGGACCGAAGCGGTAAAGAGGGGATCGAGGCCAATCTAG 1329549

QY 421 TTTCCCATCAACCATGTAGCATCCCGCAATGAGGGGTTGCAATGGCCAAAGTGGCGCAT 480
Db 1329548 TTTCCCATCAACCATGTAGCATCCCGCAATGAGGGGTTGCAATGGCCAAAGTGGCGCAT 1329489
QY 481 GGTTCCAAAGTCTTACTTCAATCCCGCAACGGGATTTAGCTTCAAGGTTACCGCTCC 540
Db 1329488 GGTTCCAAAGTCTTACTTCAATCCCGCAACGGGATTTAGCTTCAAGGTTACCGCTCC 1329429
QY 541 TAAACATCTCCAGCGCAGCAAGGATAATGTGTGGCTTTCATCTTCAAGCGCAGCGT 600
Db 1329428 TAAACATCTCCAGCGCAGCAAGGATAATGTGTGGCTTTCATCTTCAAGCGCAGCGT 1329369
QY 601 GAGCGTTGCTCCACCCCAAGAAGCTTACCTCGTTGAACAACGGGAGGAAACCATGTGGATAG 660
Db 1329368 GAGCGTTGCTCCACCCCAAGAAGCTTACCTCGTTGAACAACGGGAGGAAACCATGTGGATAG 1329309
QY 661 CGAATCTGCGTTGATGGCGATGGTTAAACGGGATTTACGAGGAGCGTCCAGATAGTTCGCG 720
Db 1329308 CGAATCTGCGTTGATGGCGATGGTTAAACGGGATTTACGAGGAGCGTCCAGATAGTTCGCG 1329249
QY 721 TTTAGTTTCTGCTTGCAGCAACCAATTTTCGCGCTGCTTGCACAAGGACCTTCACCGCG 780
Db 1329248 TTTAGTTTCTGCTTGCAGCAACCAATTTTCGCGCTGCTTGCACAAGGACCTTCACCGCG 1329189
QY 781 TTCGTTGCTTTGGCGGTTGGGTCGCGATACCAACAACCTCGACCCACCGTATGCTCGAG 840
Db 1329188 TTCGTTGCTTTGGCGGTTGGGTCGCGATACCAACAACCTCGACCCACCGTATGCTCGAG 1329129
QY 841 AGCTTTAAACCGCTGACTCACCGCGCGGAGGAAATGGAAGGCTTAAGGAGGCGCTTC 900
Db 1329128 AGCTTTAAACCGCTGACTCACCGCGCGGAGGAAATGGAAGGCTTAAGGAGGCGCTTC 1329069
QY 901 GAAGCTGCTTCAATCAATGATTGAGAGCAAAAGTGTCCAGTTGAATGGGTTTCATGAAGCT 960
Db 1329068 GAAGCTGCTTCAATCAATGATTGAGAGCAAAAGTGTCCAGTTGAATGGGTTTCATGAAGCT 1329009
QY 961 ATATTAAACCATGTTAAGAACCAATCAATTTACTTAAGTACTTCCATAGTCAAGTGGT 1020
Db 1329008 ATATTAAACCATGTTAAGAACCAATCAATTTACTTAAGTACTTCCATAGTCAAGTGGT 1328949
QY 1021 GATCATGGAATCTTCAATCAAGTCTGCTTTGGGGCCAGTCTTTACTCTCCATCGG 1080
Db 1328948 GATCATGGAATCTTCAATCAAGTCTGCTTTGGGGCCAGTCTTTACTCTCCATCGG 1328889
QY 1081 ACCGCAGAAATGACTGCTGATTAAACAAGGAATTAAGCGGAGGACTCAATTCGCGTTCT 1140
Db 1328888 ACCGCAGAAATGACTGCTGATTAAACAAGGAATTAAGCGGAGGACTCAATTCGCGTTCT 1328829
QY 1141 TCTCGTGTGTTAAATTTCTGAGCTCTTTTGTTCATCGCGGACCTTTGGCGGTTGATCT 1200
Db 1328828 TCTCGTGTGTTAAATTTCTGAGCTCTTTTGTTCATCGCGGACCTTTGGCGGTTGATCT 1328769
QY 1201 TTTGTCCAATCGCGCGCATGCTGCTCGATATTATGCGCTGGGCTGGCATCGCTTACCT 1260
Db 1328768 TTTGTCCAATCGCGCGCATGCTGCTCGATATTATGCGCTGGGCTGGCATCGCTTACCT 1328709
QY 1261 GTTATGTTTGGCGTCAATCGCAGGAAAGACGCCATGACAAACAAGGTGAAGCGGCCACA 1320
Db 1328708 GTTATGTTTGGCGTCAATCGCAGGAAAGACGCCATGACAAACAAGGTGAAGCGGCCACA 1328649
QY 1321 GATCATTAAGAAACAGAACCAACCGTGGCCGATGACAGCTTTTGGCGGTTTCGCGCT 1380
Db 1328648 GATCATTAAGAAACAGAACCAACCGTGGCCGATGACAGCTTTTGGCGGTTTCGCGCT 1328589
QY 1381 GGCCACTGACACCGCAACCGGCTGGGCTGGAGGTGAGCGTTCGATTAAGCAGCGGTTTG 1440
Db 1328588 GGCCACTGACACCGCAACCGGCTGGGCTGGAGGTGAGCGTTCGATTAAGCAGCGGTTTG 1328529
QY 1441 GGTAAAGCCCATGTTGATGCGCAATCGTCTGAACCTGTTGAACCCGATCGGTATTGGA 1500
Db 1328528 GGTAAAGCCCATGTTGATGCGCAATCGTCTGAACCTGTTGAACCCGATCGGTATTGGA 1328469
QY 1501 CGCGTTTGTGTTTATCGCGGCGTCTGGCGCGCAATACGCGGACACCCGACGCGTGGATTT 1560

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
 TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS

Db	1328468	CGCGTTGTGTTTATCGCGCGGCTCGGCGCGCAATACGCGCAACCGGACGGTGGATTTT	1328409
Qy	1561	CGCGGCTGCGCGGTTCCGGCGAAGCTGATCTGGTCCCGCTGGTGGGTTTTCGGCGCAGC	1620
Db	1328408	CGCGGCTGCGCGGTTTCGGCGAAGCCCTGATCTGGTCCCGCTGGTGGGTTTCGGCGCAGC	1328349
Qy	1621	AGCATTTGTCACGCCCGCTGTCCAGSCCCCAAGGTGTGGCGTGGATCAACGTCGTCGTGGC	1680
Db	1328348	AGCATTTGTCACGCCCGCTGTCCAGSCCCCAAGGTGTGGCGTGGATCAACGTCGTCGTGGC	1328289
Qy	1681	AGTTGTGATGACGGCAITGGCCATCAATCTGATGTTGATGGGTAGTTTTCGGGGTTTT	1740
Db	1328288	AGTTGTGATGACGGCAITGGCCATCAATCTGATGTTGATGGGTAGTTTTCGGGGTTTT	1328229
Qy	1741	GGAATCGGTGGGCTTCGCCCAATGTTGATCGCGCGCTCGTGGAAATCTCATGATCGC	1800
Db	1328228	GGAATCGGTGGGCTTCGCCCAATGTTGATCGCGCGCTCGTGGAAATCTCATGATCGC	1328169
Qy	1801	TTCCAACTCGGGCTGAGAAACTCCAAAGTTGTGAGTGAATCAAGCTGTTGTCCAGCTG	1860
Db	1328168	TTCCAACTCGGGCTGAGAAACTCCAAAGTTGTGAGTGAATCAAGCTGTTGTCCAGCTG	1328109
Qy	1861	CTCAACTGACGAAGACCAATCAATGACCTGGTCACGGTATCGCGCCGTACTCTCCCTG	1920
Db	1328108	CTCAACTGACGAAGACCAATCAATGACCTGGTCACGGTATCGCGCCGTACTCTCCCTG	1328049
Qy	1921	CTCGCGCAGCACCCATGCAAGCGCATCTCGCGAGTCACTGCCCCGGTTCCTGGCGAT	1980
Db	1328048	CTCGCGCAGCACCCATGCAAGCGCATCTCGCGAGTCACTGCCCCGGTTCCTGGCGAT	1327989
Qy	1981	GTCAATGAGCTTCGGCACCATCAATATTTGTTTCAGTTTCAACATGCGCTCAGACAGGGA	2040
Db	1327988	GTCAATGAGCTTCGGCACCATCAATATTTGTTTCAGTTTCAACATGCGCTCAGACAGGGA	1327929
Qy	2041	CTTACCCCTGGCTGGCGCGGAAACCTCTGGAATTCATCGAGATATTTGTCGTGAGCAG	2100
Db	1327928	CTTACCCCTGGCTGGCGCGGAAACCTCTGGAATTCATCGAGATATTTGTCGTGAGCAG	1327869
Qy	2101	GCCCTGCCAAGTGGTGAAGAAGTAAGCAAGCAAGCCATTGTTGGCAGCTGACTGCAA	2160
Db	1327868	GCCCTGCCAAGTGGTGAAGAAGTAAGCAAGCAAGCCATTGTTGGCAGCTGACTGCAA	1327809
Qy	2161	CAAGTTTCTACCGTCTACGCGCGGTTCTCCACCCACGATTAATGATGGAATAGCTGG	2220
Db	1327808	CAAGTTTCTACCGTCTACGCGCGGTTCTCCACCCACGATTAATGATGGAATAGCTGG	1327749
Qy	2221	CTGATGAATTCAGAAGCGGGCAGCCCTCTCCGCCATGAACTCAGCGCGCTCCGCTGAG	2280
Db	1327748	CTGATGAATTCAGAAGCGGGCAGCCCTCTCCGCCATGAACTCAGCGCGCTCCGCTGAG	1327689
Qy	2281	CTCTGGACCGTAGGAAGAAATACCAACGTAAAGAGCTTTCCAGACGCAACAATGTACG	2340
Db	1327688	CTCTGGACCGTAGGAAGAAATACCAACGTAAAGAGCTTTCCAGACGCAACAATGTACG	1327629
Qy	2341	CAATGCGTACATGGTTTCTTCCAAAGGATATCT	2374
Db	1327628	CAATGCGTACATGGTTTCTTCCAAAGGATATCT	1327595

RESULT 3

US-09-746-660A-47/C
; Sequence 47, Application US/09746660A
; Publication No. US20030049804A1
GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon

Case	Match	41 88. Score 993:	DB 10: Length 993;
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1
12	1	1	1
13	1	1	1
14	1	1	1
15	1	1	1
16	1	1	1
17	1	1	1
18	1	1	1
19	1	1	1
20	1	1	1
21	1	1	1
22	1	1	1
23	1	1	1
24	1	1	1
25	1	1	1
26	1	1	1
27	1	1	1
28	1	1	1
29	1	1	1
30	1	1	1
31	1	1	1
32	1	1	1
33	1	1	1
34	1	1	1
35	1	1	1
36	1	1	1
37	1	1	1
38	1	1	1
39	1	1	1
40	1	1	1
41	1	1	1
42	1	1	1
43	1	1	1
44	1	1	1
45	1	1	1
46	1	1	1
47	1	1	1
48	1	1	1
49	1	1	1
50	1	1	1
51	1	1	1
52	1	1	1
53	1	1	1
54	1	1	1
55	1	1	1
56	1	1	1
57	1	1	1
58	1	1	1
59	1	1	1
60	1	1	1
61	1	1	1
62	1	1	1
63	1	1	1
64	1	1	1
65	1	1	1
66	1	1	1
67	1	1	1
68	1	1	1
69	1	1	1
70	1	1	1
71	1	1	1
72	1	1	1
73	1	1	1
74	1	1	1
75	1	1	1
76	1	1	1
77	1	1	1
78	1	1	1
79	1	1	1
80	1	1	1
81	1	1	1
82	1	1	1
83	1	1	1
84	1	1	1
85	1	1	1
86	1	1	1
87	1	1	1
88	1	1	1
89	1	1	1
90	1	1	1
91	1	1	1
92	1	1	1
93	1	1	1
94	1	1	1
95	1	1	1
96	1	1	1
97	1	1	1
98	1	1	1
99	1	1	1
100	1	1	1

	Query Match 41.8%; Score 993; DB 10; Length 993; Best Local Similarity 100.0%; Pred. No. 0; Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	62 CTGAACCTTTTCAGAGTAACCTTAGGCCGCAATCCCTCGATCTCTGCAATCAACAGCGCG 121
DB	993 CTGAACCTTTTCAGAGTAACCTTAGGCCGCAATCCCTCGATCTCTGCAATCAACAGCGCG 934
QY	122 TCTGTGAGTCTAGCTAGACATCTAGATTCCAGGCCCATCGTTGCCAATCATCGGTGTG 181
DB	993 TCTGTGAGTCTAGCTAGACATCTAGATTCCAGGCCCATCGTTGCCAATCATCGGTGTG 874
QY	182 TCAATGGGTATCTCATCGAGGAGGATCACTTCTCTGCTTTTAGCATGGGAGCAGCTCG 241
DB	873 TCAATGGGTATCTCATCGAGGAGGATCACTTCTCTGCTTTTAGCATGGGAGCAGCTCG 814
QY	242 GTTTTCGGGAGAAGTCCCAACCAAGGCGCTCGCGCAATTGCTTCACAAAACTTCCTCGCG 301
DB	813 GTTTTCGGGAGAAGTCCCAACCAAGGCGCTCGCGCAANTGCTTCACAAAACCTTCCTCGCG 754
QY	302 GACGGGCAATGAGATAGCGGCGCTCGCGCCCAACAGGACCAATCGAGCGCGCGTCCAGGTCA 361
DB	753 GACGGGCAATGAGATAGCGGCGCTCGCGCCCAACAGGACCAATCGAGCGCGCGTCCAGGTCA 694
QY	362 CGGTCTTGGAAGCAATCTTTTGGGACGAGGGTAAGACGGGCAATCGGACGCCCAATCTAGT 421
DB	693 CGGTCTTGGAAGCAATCTTTTGGGACGAGGGTAAGACGGGCAATCGGACGCCCAATCTAGT 634
QY	422 TTCCCATCAACATGTAGGCAATCCCGCAATGAGGGGTTGCCAATGGCCAAAGTGGCGCATG 481
DB	633 TTCCCATCAACATGTAGGCAATCCCGCAATGAGGGGTTGCCAATGGCCAAAGTGGCGCATG 574
QY	482 GTTCCAAAGTTCTACTACTTCAATCCCGCAACGGGATTAGCTTCA CGGTTACCGCTCCT 541
DB	573 GTTCCAAAGTTCTACTACTTCAATCCCGCAACGGGATTAGCTTCA CGGTTACCGCTCCT 514
QY	542 AAAACATCTCCACCGCGCAAGAGGTAATGTGTGGCTTCATCTTCCAGCGCAGCGGTG 601
DB	513 AAAACATCTCCACCGCGCGCAAGAGGTAATGTGTGGCTTCATCTTCCAGCGCAGCGGTG 454
QY	602 ACGTTCGTCCACCCCAAGAGACTACTCTGTTGAAACCGGAGGAAACCATGTGGATAGC 661

; CURRENT FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 09/606740
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 09/603124
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/142101
 ; PRIOR FILING DATE: 1999-07-02
 ; PRIOR APPLICATION NUMBER: 60/148613
 ; PRIOR FILING DATE: 1999-08-12
 ; PRIOR APPLICATION NUMBER: 60/187970
 ; PRIOR FILING DATE: 2000-03-09
 ; PRIOR APPLICATION NUMBER: DE 19931420.9
 ; PRIOR FILING DATE: 1999-07-08
 ; NUMBER OF SEQ ID NOS: 125
 ; SOFTWARE: PatentIn Vers. 2.0
 ; SEQ ID NO 51
 ; LENGTH: 822
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(799)
 ; OTHER INFORMATION: RXA01394
 ;
 ; US-09-746-660A-51
 ;
 ; Query Match 34.6%; Score 822; DB 10; Length 822;
 ; Best Local Similarity 100.0%; Pred. No. 7.1e-264;
 ; Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ;
 QY 925 GACAAAGTGTCCAGTTCATGAGGGTTCATGAGCTATATAAACCAATGTAAGAACCAA 984
 DB 1 GAGCAAGGTGCCAGTTCATGAGGGTTCATGAGCTATATAAACCAATGTAAGAACCAA 60
 ;
 QY 985 TCATTTTTAACTTCCATAGGTCAAGATGATCAATGAGAAATCTTATTACAGG 1044
 DB 61 TCATTTTTAACTTCCATAGGTCAAGATGATCAATGAGAAATCTTATTACAGG 120
 ;
 QY 1045 TCTGCTTTTGGGGCGAGTCTTTTACTGTCCATGAGCCGAGATGATCTGTGATTA 1104
 DB 121 TCTGCTTTTGGGGCGAGTCTTTTACTGTCCATGAGCCGAGATGATCTGTGATTA 180
 ;
 QY 1105 ACAAGAAATAGCGGAGGACTCATTTGCGGTTCTCTCGTGTGTTAAATTTCTGAGCT 1164
 DB 181 ACAAGAAATAGCGGAGGACTCATTTGCGGTTCTCTCGTGTGTTAAATTTCTGAGCT 240
 ;
 QY 1165 CTTTGTGTTTATCGCGGCACTTGGGGGTGATCTTTTGTCCAAATGCGCGCGGATGCT 1224
 DB 241 CTTTGTGTTTATCGCGGCACTTGGGGGTGATCTTTTGTCCAAATGCGCGCGGATGCT 300
 ;
 QY 1225 GCTCGATATATGCGTGGGGTGGATCGCTTACTGTATGTTTGGCGTATGCGAGC 1284
 DB 301 GCTCGATATATGCGTGGGGTGGATCGCTTACTGTATGTTTGGCGTATGCGAGC 360
 ;
 QY 1285 GAAGAGCCATGACAAACAAAGGTGGAAGCGCACAGATCAITGAAGAAACAGAACCAAC 1344
 DB 361 GAAGAGCCATGACAAACAAAGGTGGAAGCGCACAGATCAITGAAGAAACAGAACCAAC 420
 ;
 QY 1345 CTTGCGCGATGACAGCCCTTTGGCGGTTTGGCGGTTGCGGCGGCGCACTGACAGCGG 1404
 DB 421 CTTGCGCGATGACAGCCCTTTGGCGGTTTGGCGGTTTGGCGGCGGCGCACTGACAGCGG 480
 ;
 QY 1405 GCGGTTGGAGGTGAGCGTGCATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAAT 1464
 DB 481 GCGGTTGGAGGTGAGCGTGCATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAAT 540
 ;
 QY 1465 CTTGCTGACCTGTTCAACCCGAATTCGATATTTGACGCGGTTTGTGTTTATCGCGCGGT 1524
 DB 541 CTTGCTGACCTGTTCAACCCGAATTCGATATTTGACGCGGTTTGTGTTTATCGCGCGGT 600
 ;
 QY 1525 CCGCGCGCAATAGCGGCAACCGGACGCGGTGGAATTTTCCGCTGCGGCGGTTTGGCGCAAG 1584
 DB 361 CCGCGCGCAATAGCGGCAACCGGACGCGGTGGAATTTTCCGCTGCGGCGGTTTGGCGCAAG 420

QY 1436 GTTGGGTAAGCCCATTTGATGGAATCGTGTGACTGTTGAACCGAATGGTAT 1495
 Db 421 GTTGGGTAAGCCCATTTGATGGAATCGTGTGACTGTTGAACCGAATGGTAT 480
 QY 1496 TTGGACGCGTTGTGTTATCGCGCGCTCGCGCGCAATAGCGGACACCGACGCGTGG 1555
 Db 481 TTGGACGCGTTGTGTTATCGCGCGCTCGCGCGCAATAGCGGACACCGACGCGTGG 540
 QY 1556 ATTTTCCGCGCTGCGCGGTTTCGCGCAAGCCTGATCTGGTTCCTGCTGGTGGGTTTCGCG 1615
 Db 541 ATTTTCCGCGCTGCGCGGTTTCGCGCAAGCCTGATCTGGTTCCTGCTGGTGGGTTTCGCG 600
 QY 1616 GCAGCAGCATTTGATGAGCGCGCTGTCAGCCCGCAAGGTGCGGCTGGATCAAGCTGTC 1675
 Db 601 GCAGCAGCATTTGATGAGCGCGCTGTCAGCCCGCAAGGTGCGGCTGGATCAAGCTGTC 660
 QY 1676 GTGGCAGTTGTCATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 1726
 Db 661 GTGGCAGTTGTCATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 711

RESULT 7
 US-09-738-626-3455
 ; Sequence 3455, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738, 626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 3455
 ; LENGTH: 708
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3455

Query Match 29.8%; Score 708; DB 9; Length 708;
 Best Local Similarity 100.0%; Pred. No. 1e-225;
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1016 ATGGTGATCATGAAATCTTCAATACAGGTCTGCTTTGGGGCCAGCTTTTACTGTCC 1075
 Db 1 ATGGTGATCATGAAATCTTCAATACAGGTCTGCTTTGGGGCCAGCTTTTACTGTCC 60
 QY 1076 ATCGGACCGAGAAATGCTGATTAACAGGAATTAAGCGGACGACTCATTTGGC 1135
 Db 61 ATCGGACCGAGAAATGCTGATTAACAGGAATTAAGCGGACGACTCATTTGGC 120
 QY 1136 GTTCTCTCGTGTTTAAATTCGAGTCTTTTGTTCATCGCGCACCTTGGCGGTT 1195
 Db 121 GTTCTCTCGTGTTTAAATTCGAGTCTTTTGTTCATCGCGCACCTTGGCGGTT 180
 QY 1196 GATCTTTTGTCCAAATCGCGCGCATGCTGCTGATATTAATGCGCTGGGCTGGCATCGCT 1255
 Db 181 GATCTTTTGTCCAAATCGCGCGCATGCTGCTGATATTAATGCGCTGGGCTGGCATCGCT 240

QY 1256 TACCTGTTATGTTTGGCTCATGCGGAAAGACGCGCATGACAAACAAAGGTGGAAGCG 1315
 Db 241 TACCTGTTATGTTTGGCTCATGCGGAAAGACGCGCATGACAAACAAAGGTGGAAGCG 300
 QY 1316 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGCGGTTTCG 1375
 Db 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGCGGTTTCG 360
 QY 1376 CGCTGGCCCATGACACCGGCAACCGGCTGGGCTGGAGGTGAGCGTGAATAGCAGCGG 1435
 Db 361 CGCTGGCCCATGACACCGGCAACCGGCTGGGCTGGAGGTGAGCGTGAATAGCAGCGG 420
 QY 1436 GTTTGGGTAAGCCCATTTGATGGAATCGTGTGACTGTTGAACCGCAATGCGGTAT 1495
 Db 421 GTTTGGGTAAGCCCATTTGATGGAATCGTGTGACTGTTGAACCGCAATGCGGTAT 480
 QY 1496 TTGGACGCGTTGTGTTATCGCGCGCTCGCGCGCAATACCGGACACCGGACGCTGG 1555
 Db 481 TTGGACGCGTTGTGTTATCGCGCGCTCGCGCGCAATACCGGACACCGGACGCTGG 540
 QY 1556 ATTTTCCGCGCTGCGCGCTTCGCGCAAGCCTGATCTGGTTCCTGCTGGTGGTTCGCG 1615
 Db 541 ATTTTCCGCGCTGCGCGCTTCGCGCAAGCCTGATCTGGTTCCTGCTGGTGGTTCGCG 600
 QY 1616 GCAGCAGCATTTGTCAGCGCGCTGTCAGCCCGCAAGGTGCGGCTGGATCAAGCTGTC 1675
 Db 601 GCAGCAGCATTTGTCAGCGCGCTGTCAGCCCGCAAGGTGCGGCTGGATCAAGCTGTC 660
 QY 1676 GTGGCAGTTGTCATGACCGCATTTGGCCATCAAACTGATGTTGATGGT 1723
 Db 661 GTGGCAGTTGTCATGACCGCATTTGGCCATCAAACTGATGTTGATGGT 708

RESULT 8
 US-10-166-142-9
 ; Sequence 9, Application US/10166142
 ; Publication No. US20030124687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GUNJI, YOSHIYA
 ; APPLICANT: YASUEDA, HISASHI
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
 ; TITLE OF INVENTION: ASSIMILATING BACTERIUM
 ; FILE REFERENCE: 223789US
 ; CURRENT APPLICATION NUMBER: US/10/166,142
 ; CURRENT FILING DATE: 2002-06-11
 ; PRIOR APPLICATION NUMBER: JP 2001-1777075
 ; PRIOR FILING DATE: 2001-06-12
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 9
 ; LENGTH: 712
 ; TYPE: DNA
 ; ORGANISM: Brevibacterium lactofermentum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(375)
 ; OTHER INFORMATION:
 US-10-166-142-9

Query Match 29.4%; Score 696.8; DB 14; Length 712;
 Best Local Similarity 99.6%; Pred. No. 5.7e-222;
 Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 1016 ATGGTGATCATGAAATCTTCAATACAGGTCTGCTTTGGGGCCAGCTTTTACTGTCC 1075
 Db 1 ATGGTGATCATGAAATCTTCAATACAGGTCTGCTTTGGGGCCAGCTTTTACTGTCC 60
 QY 1076 ATCGGACCGAGAAATGCTGCTGATTAACAGGAATTAAGCGGACGACTCATTTGGC 1135
 Db 61 ATCGGACCGAGAAATGCTGCTGATTAACAGGAATTAAGCGGACGACTCATTTGGC 120
 QY 1136 GTTCTCTCGTGTTTAAATTCGAGTCTTTTGTTCATCGCGCACCTTGGCGGTT 1195

Query Match	26.4%;	Score 627;	DB 9;	Length 627;
Best Local Similarity	100.0%;	Pred. No. 1.3e-198;		
Matches 627;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1726	GTTTTCGGGGTTTTGGAAATCGGTGGCCCTTCGCCAAATGTTGATGCGCGCGPGFGGA	1785	
DB	627	GTTTTCGGGGTTTTGGAAATCGGTGGCCCTTCGCCAAATGTTGATGCGCGCGTCGTGGGA	568	
QY	1786	AATCTCATGATGGCTTCCAATCTGGCGTCAAGAAATCCCAAGTTGTTGAGTCGAATCAAG	1845	
DB	567	AATCTCATGATGGCTTCCAATCTGGCGTCAAGAAATCTCCAAGTTGTTGAGTCGAATCAAG	508	
QY	1846	GCCTGTTCGACGTGCTCAAATGCAGGAAGCAACCAATCAATGCACTGGTTCACGGTATCCGC	1905	
DB	507	GCCTGTTCGACGTGCTCAAATGCAGGAAGCAACCAATCAATGCACTGGTTCACGGTATCCGC	448	
QY	1906	GCCGTACTCTCTTGTCTGGCGAGACCCCATGCAAGCGCATCTCGCGAAGTCAGTCGCC	1965	
DB	447	GCCGTACTCTCTTGTCTGGCGAGACCCCATGCAAGCGCATCTCGCGAAGTCAGTCGCC	388	
QY	1966	GGTTTCCTGGGCGATGATTAATGAGCTTCGGAGACCATCAATATGTTTCAGTTCACAT	2025	
DB	387	GGTTTCCTGGGCGATGATTAATGAGCTTCGGAGACCATCAATATGTTTCAGTTCACAT	328	
QY	2026	GCCCTCAGACAGAGGACTTACCCCTGGCTGGCGGGGAAACCCCTCTGGAAATTCATCGAGATA	2085	
DB	327	GCCCTCAGACAGAGGACTTACCCCTGGCTGGCGGGGAAACCCCTCTGGAAATTCATCGAGATA	268	
QY	2086	TTTGTCCGTGAGCAGCCCTCGGCAGTGGTCAGAAAGCAATGACGCCAAGCAATGTT	2145	
DB	267	TTTGTCCGTGAGCAGCCCTCGGCAGTGGTCAGAAAGCAATGACGCCAAGCAATGTT	208	
QY	2146	GGCAGCTGACTGCAGAACAAAGTTCTCACCGTCATCGCCCGGTTCTCCACCCAAAGATTAAT	2205	
DB	207	GGCAGCTGACTGCAGAACAAAGTTCTCACCGTCATCGCCCGGTTCTCCACCCAAAGATTAAT	148	
QY	2206	GATGGAATAGCTTGGCTGATGAATCAGAGAGGGCGAGCCCTCTCGCGCATGAACCTCAGC	2265	
DB	147	GATGGAATAGCTTGGCTGATGAATCAGAGAGGGCGAGCCCTCTCTCGCGCATGAACCTCAGC	88	
QY	2266	CGCCTCCGCTGTGAGCTCTTGAGACCGTAGGAAGAAATACCCAGTAAGAGCCTTTCAGA	2325	
DB	87	CGCCTCCGCTGTGAGCTCTTGAGACCGTAGGAAGAAATACCCAGTAAGAGCCTTTCAGA	28	
QY	2326	CGCAACAAATGTCAGCAATGGTACAT	2352	
DB	27	CGCAACAAATGTCAGCAATGGTACAT	1	

RESULT 10
US-10-627-476-669/c
; Sequence 669, Application US/10627476
; Publication No. US20040030116A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CPCN
; CURRENT APPLICATION NUMBER: US/10/627,476
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 09/602,787
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US9N 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08

RESULT 11
US-10-282-122A-17521/c
; Sequence 17521, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

Query Match	12.9%;	Score 305.8;	DB 12;	Length 1095;
Best Local Similarity	67.1%;	Pred. No. Be-91;		
Matches 433;	Conservative 0;	Mismatches 212;	Indels 0;	Gaps 0;
Qy	1730	TCGCGGGTTTTGAATCGGTGGCTTCGCCCAATGTTCATGCGCGGTGTGGGAAATC	1789	
Db	1094	TGTTGAACCTTCAGGGCAGTAGACCGGCCCAATATTAATTCAGCGTCTTGGCAACA	1035	
Qy	1790	TCATCGATCGCTTCCAACTCGCGGTGAGAAAACTCCAAGTGTTCAGTGAATCAAGGCTG	1849	
Db	1034	GTGTCAATATAATTGGTTCCTCAACGCTGAACCTCAAGATTATTAAGAGCTCCAGATT	975	
Qy	1850	TTGTCCAGCTGCTCAACTGACGAGCACCAAATCAATCACTGTCACGGTATCCGCGCG	1909	
Db	974	TGATCAACTGCGCTACCGAGCATGCTCCGATCAACGACTGGTTACCGTTGGGCCCA	915	
Qy	1910	TACTCTCCTTGTCTCGCGCAGCACCCATGCGAAGCGCCATCTGCGCAAGTGACTCCCGGT	1969	
Db	914	TAATCGCTTGTTCGCGCAGCACCCAGCAATAGCAGCTGTGCCAAGTTGGCCACGT	855	
Qy	1970	TCTGGGGAGTGATTGAGCTTGGGACCATATCAATATTTTCAGTTCACATGTGCC	2029	
Db	854	CGTAACGGCTAAATCATTCAGTGCACGCAATGTCAAGATTTTTTAGCGTTGAGCATCT	795	
Qy	2030	TCAGACGGAGCTTACCTTGCCTGGCGCGGCAACCTCTCGAAATTCATCGAGATTGT	2089	
Db	794	TTGCCCAAGACTTTCCTGTCTGCTGCCGTGAATCCATGGCAGCGCCTCGAGATACCT	735	
Qy	2090	TCCGTGAGCAGGCCCTTGGCGCAAGTGCTGAGAAAGCAATGACGCCAAGACCATTTGTGCA	2149	
Db	734	TCAGTTAAAGACCCCTGTGCAAGTGGCGAAATGCAATAACACCCAGTTCGGTATTTCGC	675	

QY 2150 GCTGACTGCAACAAAGTTCTCACCGTCGCGCGGTTCTCCACCCCAAGTAATGATG 2209
Db 674 GCAGACTCTAACAGGTTCTCGCCATCTTCACTTGGTTCTTACCCAGCGGTTCAAGATA 615
QY 2210 GAATAGCTTGGTGATCAATACAGAGCGGCGAGCGCTCTCCGCGCATGAACTCAGCGCGC 2269
Db 614 GAATAGCTCGGTTGATGAATCAGAGCGGCGAGCGCTCTCCAGCTAGATTTCCGAGATC 555
QY 2270 TCCGCTGTGAGCTCTTGACCGTAGGAGAAATACCCACGTAAGAGCCTTTCCAGACGA 2329
Db 554 TCTGCGGTTAGTTCAAGTCCATAGGAGAAATTCACATAGCAGCGCTTGGCGGATGCC 495
QY 2330 ACAAATGACCAATGCGTACATGTTCTTCCAAAGGAGTATCT 2374
Db 494 ACGATGTCTCGAAAGCATATGCAAGTTCTTCCAAAGGCGTATCT 450

RESULT 12
US-10-282-122A-17523/c
; Sequence 17523, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17523
; LENGTH: 879
; TYPE: DNA
; ORGANISM: *Corynebacterium diptheriae*
US-10-282-122A-17523

Query Match 12.8%; Score 302.8; DB 12; Length 879;
Best local similarity 59.2%; Pred. No. 7e-90;
Matches 517; Conservative 0; Mismatches 357; Indels 0; Gaps 0;
QY 81 ACTAAGCGCGCAATCCCTCGATTGCTGCATCAACGAGCGGCTGTGAGTCTAGTAGAG 140

Db 874 ACAGTGAATGATGTCCTCGAGGCGTGCATCCACTACCGCTTGAAGTCAACCGCTCTAGGG 815
QY 141 ATCTAGATTCAGAGCGGCATCGTTGCCAATACATCGGTGTGTCAATGGGTATCTCATCGA 200
Db 814 CTGAGATTTCCAAACCGCATCGCTGCCAATACAGCGAATTCGTCGACCTCGCTATCTA 755
QY 201 GGAGGATCACTTCTCTCTGCTTTTATGATGGAGAGCGCTTGGGTTTCGGGAAGAGTCCCC 260
Db 754 GCATCACTACCTCCACAGAAATCCAAAGCGGTAAAGCTTGAATATCGGGGAGCAATGCC 695
QY 261 AACCAAGGCTCGGCAATTCCTTACCAAAACCTTCGCGCGAGCGGCAATGATACGC 320
Db 694 AGCCCAATCCCAACAGCGCGCTTCCATGAATGCTCCGATGAAGGATCTCATTAACAC 635
QY 321 GCCTGGCGCCCAACAGGACCATCGACGCGCGCTCCAGGTCAAGCTTTGAAAGCAATCTT 380
Db 634 GTCGCCCTGCGGAACTTCCACCAAGAGCGCTTAATATCAAGCTTCTTGAAGCCCATCCC 575
QY 381 TGGGACGAAAGCGTAAGACGGGCAATCGAGCCCAATCTAGTTTCCCAATCAACCATGAGG 440
Db 574 GAGGGCCCAATCGCAATATCTGGCATCCGCTCCCAATCAACAGACCGCTTATGGTGAGC 515
QY 441 CATCCCGCAATGAGGGGCTTGCATATGCGCATGCTTCCAGTGTTCAGAGTTCTACTACTT 500
Db 514 GATCCCAACAGCTTTCGGAATTCGCACTGCAAAATATCGAATGACCTTAACGCTATCGAGT 455
QY 501 CACATCCCGCCCAACAGGATAGCTTTCAGGGTTACCGCTTCTTAAAAATCTCCAGCGCGCA 560
Db 454 CGCAACCAAGATACAGGTGTCTTTTCCCGGTTACAGCGCCAAAGCAATCTCCGCGAGCA 395
QY 561 GCAAGGATAATGTGTGCGCTTCACTTCCAGCGGAGCGGTGCGGTGCTCCACCCCAAG 620
Db 394 ACAACGCAACAGAAATGCGACTCTCTCGATACGATATGCAATGTGGGCTTATCCATCT 335
QY 621 AGCTTACTCTGTTGACACAGGAGGAAACCATGTGATAGCAATCTGCGTTGATGCGCA 680
Db 334 CGCGCCACACTTTCGCAATCCCGCGGCAAAACCATGTGCGCAAGATCTGCGTTGACAGTA 275
QY 681 TGGTTTAAACGGGATTTTCCAGCAAGCGCTCCAGATAGTTTGGCTTTAGTTTCTGCTTCAGCA 740
Db 274 CAGCAATGGAATCCGCTCAATTTCTCCCGCAACTGAGCGTTTGTTCGGCTTCAGAA 215
QY 741 ACACATTTTCCGGCTGTGTCACAGAGCACTTCCACCGCTTCCGTTTGGCGCGGT 800
Db 214 GTGCCATACGTCTGCGGATTCACAGTACCTTCGCGCGCATCGATGTCAGTGCTG 155
QY 801 GGTGCGCGATACCAACACTCGACCCAGTGTGCTCCAGAGCTTTAAACCGCTGACTCA 860
Db 154 AGTACGGCGTACCACTACTCGACAGTCTCCCGCTCCACGCTTAATCTCTGACTTA 95
QY 861 CCGCGGAGGGGAAATGGAAGGGCTTAAGAGGCGCTTCCAGGCGCTTCCAGTCCCTCATCAATGA 920
Db 94 CTGCTGAAGAGGAAATTCCTTAATGCAAAAGCGCATCTTCGAAAGCTGCTTCTGCGATTA 35
QY 921 TTGAGAGCAAGTGTCCAGTTCGAATGGGTTTCAAT 954
Db 34 TTGCGAGCAAGGTTTCTTAATGAAGCGGATTCAT 1

RESULT 13
US-10-369-493-35374/c
; Sequence 35374, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35374
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-35374

Query Match 6.9%; Score 164.4; DB 15; Length 1026;
Best Local Similarity 57.2%; Pred. No. 1.8e-43;
Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;
QY 1745 TCGGTGGCTTCGCGCAAAATGTTGATGCGCGGTCTGGGAATCTCATCGATCGCTCC 1804
DB 1022 TCGGCAGATTTTGGCCAGAGTTGATATCCGATCTTGGCGTAACGGTGGATTTGGGCC 963
QY 1805 AACTCGCGCTCAGAAAATCTCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGTCTCA 1864
DB 962 AGCTCTCGGTAGAGAACTCGCATATCGAGTGTCTTTCAGCGAGTCTCTCGACCTGTTCG 903
QY 1865 ACTGACGAAGCACCAATCAATGCACTGTGTCAGGTATCCGCGCGTACTCTCTTGTCTCG 1924
DB 902 ACACGGCTTGGCCCAATCAATGCTGAGGTAATGCGGCCGCG-----CG 861
QY 1925 CGCAGCACCATGCAAGCGGCATCTCGCAAGTGAATGCTCGCGGTCTCTTGGCGGATGTCA 1984
DB 860 CGCAGAACCCGCGCAATTTGCCATCTGCGCAGCGTCTGCGCAGCGTCTGCGCAATGCTG 801
QY 1985 TTGAGCTTGGGACCAATCAATATTTTCAAGTTTCAATGCTTCAACATGCTTCAAGAGGATTA 2044
DB 800 TTCAGCGCGGAGTGTCTTTCGACATTTGCGCTGTTGAGGAAGCGCGGTTGAGTGACTTG 741
QY 2045 CCCTGGTGGCGGGAACCTCTGGAATTCATCGAGATTTTGTTCGTGAGCAGGCGCC 2104
DB 740 CTCTGTGAGGACGCGTCTGCCATCGGCACACCGCCAGATATTTGCTGTCAGCATGCC 681
QY 2105 TGCGCAAGTGTGAGAAAGCAATGACCGCAAGACCATTTGTTGGCAGCTGACTGCAACAG 2164
DB 680 TGCGCCAGCGCGGAAGAGATGAGAGCGGATACCCAGTCTTTCACAGGTAT----- 629
QY 2165 TTCTCACCGTCTACCGCGGTCTCTCACCCAAAGATTAATGATGAATAGCTTGGCTGA 2224
DB 628 -----CGACAGACCGTCTTCTCGATCCAGCGTGTATCATCGATAGCTCGGTG 576
QY 2225 TGAATCAGAGCGGCGAGCCCTCTCGGCATGAACTCAGCGCCCTCGCTGTGAGTCT 2284
DB 575 TGGATGATGCGAGCGGTCTCGAGATCTTTTCAAGATAGCGCGCTCGCGGCTCGCGTTC 516
QY 2285 GGACCGTAGGAAGAAATACCCAGTAAGAGCTTTTCCAGACGCAACATGTCAACGAAT 2344
DB 515 GAGTTGTAGAGAGATGCCAGATAGAGCGCTTGGCGAGCGCAGCATCTGTCTCAGC 456
QY 2345 GCGTACATGTTTCTTCCAAAGGAGTAT 2372
DB 455 GCGCGCGAGGTTTCTCTCAAGCGGTGTGT 428

RESULT 14
US-10-369-493-38826/c
; Sequence 38826, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38826
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38826

Query Match 6.9%; Score 164.4; DB 15; Length 1026;
Best Local Similarity 57.2%; Pred. No. 1.8e-43;
Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;
QY 1745 TCGGTGGCTTCGCGCAAAATGTTGATGCGCGGTCTGGGAATCTCATCGATCGCTCC 1804
DB 1022 TCGGCAGATTTTGGCCAGAGTTGATATCCGATCTTGGCGTAACGGTGGATTTGGGCC 963
QY 1805 AACTCGCGCTCAGAAAATCTCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGTCTCA 1864
DB 962 AGCTCTCGGTAGAGAACTCGCATATCGAGTGTCTTTCAGCGAGTCTCTCGACCTGTTCG 903
QY 1865 ACTGACGAAGCACCAATCAATGCACTGTGTCAGGTATCCGCGCGTACTCTCTTGTCTCG 1924
DB 902 ACACGGCTTGGCCCAATCAATGCTGAGGTAATGCGGCCGCG-----CG 861
QY 1925 CGCAGCACCATGCAAGCGGCATCTCGCAAGTGAATGCTCGCGGTCTCTTGGCGGATGTCA 1984
DB 860 CGCAGAACCCGCGCAATTTGCCATCTGCGCAGCGTCTGCGCAGCGTCTGCGCAATGCTG 801
QY 1985 TTGAGCTTGGGACCAATCAATATTTTCAAGTTTCAATGCTTCAACATGCTTCAAGAGGATTA 2044
DB 800 TTCAGCGCGGAGTGTCTTTCGACATTTGCGCTGTTGAGGAAGCGCGGTTGAGTGACTTG 741
QY 2045 CCCTGGTGGCGGGAACCTCTGGAATTCATCGAGATTTTGTTCGTGAGCAGGCGCC 2104
DB 740 CTCTGTGAGGACGCGTCTGCCATCGGCACACCGCCAGATATTTGCTGTCAGCATGCC 681
QY 2105 TGCGCAAGTGTGAGAAAGCAATGACCGCAAGACCATTTGTTGGCAGCTGACTGCAACAG 2164
DB 680 TGCGCCAGCGCGGAAGAGATGAGAGCGGATACCCAGTCTTTCACAGGTAT----- 629
QY 2165 TTCTCACCGTCTACCGCGGTCTCTCACCCAAAGATTAATGATGAATAGCTTGGCTGA 2224
DB 628 -----CGACAGACCGTCTTCTCGATCCAGCGTGTATCATCGATAGCTCGGTG 576
QY 2225 TGAATCAGAGCGGCGAGCCCTCTCGGCATGAACTCAGCGCCCTCGCTGTGAGTCT 2284
DB 575 TGGATGATGCGAGCGGTCTCGAGATCTTTTCAAGATAGCGCGCTCGCGGCTCGCGTTC 516
QY 2285 GGACCGTAGGAAGAAATACCCAGTAAGAGCTTTTCCAGACGCAACATGTCAACGAAT 2344
DB 515 GAGTTGTAGAGAGATGCCAGATAGAGCGCTTGGCGAGCGCAGCATCTGTCTCAGC 456
QY 2345 GCGTACATGTTTCTTCCAAAGGAGTAT 2372
DB 455 GCGCGCGAGGTTTCTCTCAAGCGGTGTGT 428

RESULT 15
US-10-369-493-38124/c
; Sequence 38124, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

Thu Mar 18 12:30:59 2004

```

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38124
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38124

Query Match          6.9%; Score 164.4; DB 15; Length 1028;
Best Local Similarity 57.2%; Pred.No. 1.8e-43;
Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;

Qy 1745 TCGGTGGCTTCGCCCAATGTTGATCGCGGCTGTTGGGAAATCTCATCGATCGCTCC 1804
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1805 AACTCGGCGTCAGAAACTCCAACTGTTGAGTGAATCAAGGCTGTTGTCAGTGTCTCA 1864
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1865 ACTGACGACACCAATCAATGCACTGCTCAGGATATCGGCGCTGCTCTCTTGTCTCG 1924
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1925 CGCAGCACCATGCAAGCGCCATCTGGCGAAGTGAATCGCGCGTTCCTTGGGCGATGCA 1984
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1985 TTGAGCTTGGGACCATATCAATATGTTTACGTTTCAATGTCCTTCAGACGAGGACTTA 2044
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2045 CCCTGGCTGGCGGGGACCTCTCGAATTCATCGAGATATTTGTCGTGACGAGGCC 2104
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2105 TCGCAAGTGTGAGAAAGCAATGACGCCAAGACCAATTTGTCGACGCTGACGCAACAG 2164
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2165 TTCTACCGTATCGCCGCTTCCTCCACCAAGATTAATGATGGAATAGCTTGGCTGA 2224
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2225 TGAATCAGAGCGGCGAGCCCTCTCCGCCATGAATCAAGCGCCCTCGCTGTGAGCTCT 2284
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2285 GGACCGTAGGAGAAATACCCAGTAAGAGCCCTTTCAGACGCAACAAATGTCAGCAAT 2344
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2345 GGTACATGTTTCTTCAAGGATAT 2372
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2372 GCGTACATGTTTCTTCAAGGATAT 2400
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2400 GCGCCGAGGTTTCTTCAAGGATAT 2428
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: March 17, 2004, 09:46:30
Job time : 870.071 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:43:00 ; Search time 6528.5 Seconds
(without alignments)
10858.975 Million cell updates/sec

Title: US-09-105-117K-1
Perfect score: 2374
Sequence: 1 ccatttgctgaaggtgttac.....ttttccaaaggagtattct 2374

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pla:*
20: em_ges_vrc:*
21: em_ges_fun:*
22: em_ges_man:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rod:*
26: em_ges_phg:*
27: em_ges_vri:*
28: gb_ges1:*
29: gb_ges2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.6	3.6	657	13	BQ155035
2	73.2	3.1	3237	28	BH770958
3	72	3.0	860	28	AF075981
4	60.4	2.5	830	29	CNS01MQH

5	56.8	2.4	1034	28	B2554294
6	54.2	2.3	897	28	B2560550
7	51.4	2.2	1620	28	B2568946
8	48.8	2.1	1101	29	CNS017SY
9	48.6	2.0	833	28	B2556963
10	48.6	2.0	925	29	CNS0091P
11	47.6	2.0	852	13	BX333687
12	46	1.9	614	14	CA830247
13	46	1.9	720	14	CD881613
14	44.4	1.9	620	14	CA829022
15	44.4	1.9	644	13	BQ704105
16	44.4	1.9	664	29	CNS03K4J
17	44.4	1.9	1201	13	BX381961
18	43.6	1.8	721	28	B2562255
19	43.2	1.8	1395	28	B2574094
20	42.8	1.8	532	10	AW927407
21	42.8	1.8	538	14	CD861975
22	42.4	1.8	699	13	BX424825
23	42.2	1.8	885	13	BX425603
24	42	1.8	457	13	BY246241
25	42	1.8	1201	13	BX381961
26	41.8	1.8	947	29	CNS077MV
27	41.8	1.8	1083	29	CNS079WN
28	41.6	1.8	432	14	CB639421
29	41.6	1.8	558	12	BM712036
30	41.6	1.8	588	13	BU220119
31	41.6	1.8	772	14	CB668778
32	41.6	1.8	803	14	CB669972
33	41.6	1.8	813	14	CB684812
34	41.6	1.8	1201	13	BX356664
35	41.6	1.8	1201	13	BX361080
36	41.2	1.7	910	29	CNS0060N
37	41	1.7	617	12	BM729055
38	41	1.7	770	10	BE898495
39	41	1.7	861	10	BE740880
40	40.8	1.7	605	13	CA127255
41	40.8	1.7	614	13	CA141725
42	40.8	1.7	619	14	CA192339
43	40.8	1.7	619	14	CA238619
44	40.8	1.7	950	11	AY104298
45	40.6	1.7	436	28	AQ405681

ALIGNMENTS

RESULT 1
BQ155035
LOCUS NF075E08IR 5', mRNA linear EST 24-APR-2002
DEFINITION NF075E08IR1067 Irradiated Medicago truncatula cDNA clone
ACCESSION BQ155035
VERSION BQ155035.1 GI:20292094
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyl;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE 1 (bases 1 to 657)
AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weiler, J.W. and May, G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula irradiated library
JOURNAL Unpublished (2001)
COMMENT Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org

Insert Length: 657 Std Error: 0.00
 Plate: 075 row: E column: 08
 Seq primer: TCACAGGAAACAGCTATGAC.
 Location/Qualifiers
 1. .657
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"
 /clone="NF075E08IR"
 /tissue_type="seedlings"
 /dev_stage="seedling"
 /clone_lib="Irradiated"
 /note="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-ZAP XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 3.6%; Score 84.6; DB 13; Length 657;
 Best Local Similarity 48.2%; Pred. No. 3.4e-11;
 Matches 237; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

464 ATGCCAGTGGCGATGTTCCAAAGTTCTACTTCCATCCCGCCACGGATAGCT 523
 Db
 114 ACGAACAGATAGTCGAGCGCCCAAGTTTATCGACAGCAACTCCGCGAGCGCTGAT 173
 524 TCACGGGTTACCGCTCTAAACATCTCCACGCCGACAGGATTAATGTGCGCTTCA 583
 Db
 174 TGAATATCTACCGCGCGACCACTTCGCGCGCGCAGACGTTCTCTGAGTGGGGTTCA 233
 584 TCTTCCAGCGAGCGTGGTGGTTCACCCCAAGAGCTACCTCGTTGAACACGGGA 643
 Db
 234 TCTTCTACCTGCAAGTTGAGCGGATAGGCGAATCAGCCACACAGGACCGCAGTGCAG 293
 644 GGAACCATCTGATAGCAATCTGCGTGTAGTGGGATGTTAAGCGGATTTACAGCAAG 703
 Db
 294 AGCAACACGTCGCCAGACTGTGCGCGTTGACCGCGATGAAAGCAGCGAGCGAGTCAA 353
 704 CGTCCAGATAGTTGCGCTTTAGTTTCTGTTGACGACACACCAATTTTCCGCGCTTGC 763
 Db
 354 CCGGTTTCTTCATCGCCCGACGCACTCTTCTTCCAGCACTCCACCTGGCGCAGCTGCC 413
 764 ACAAGACTTCACCGCTTCGGTGTCTTTCGCGGTGGTGGCGGATACCAACTCGA 823
 Db
 414 AGCAGTTTCTTGGCTTTGTTTCGTCGCGCGCGGTACGGTACGCAACACAGCGCGCTGC 473
 824 CCACGTCATGCTCGAGAGCTTTAAAGCGCTGACTCACCGCGCGGGGAAATGMAAGG 883
 Db
 474 CCGAACATATTTTCCAGTTGCTTAATGGCTGTGAGACGCGTGTATGTATGACACAGC 533
 884 GCTAAGAGGCGGCTTCGAAGCTGCTTCAATGATGATGAGACAAAGTTCAGTTGA 943
 Db
 534 TTTTGTGCGCGGCTCAATCTCTGTTACGATATCACCGCATCCAGTGCCTGTATGTT 593
 944 ATGGGTTTCATG 955
 Db
 594 CTGTAGTCCNG 605

RESULT 2

BH770958 3237 bp DNA linear GSS 01-MAY-2002
 LOCUS
 DEFINITION LLMGtag686 MG1363 Random Sequence Tag Library Lactococcus lactis

FEATURES

subsp. cremoris genomic, genomic survey sequence.
 BH770958
 VERSION BH770958.1 GI:20373915
 KEYWORDS
 SOURCE
 ORGANISM
 Lactococcus lactis subsp. cremoris
 Lactococcus lactis subsp. cremoris
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Lactococcus.
 REFERENCE
 1. (bases 1 to 3237)
 Bolotin, A., Ehrlich, S.D. and Sorokin, A.
 Studies of genomes of dairy bacteria Lactococcus lactis
 Sci. Aliments (2002) In press
 Contact: Sorokin A
 Genetique Microbiome
 INRA
 CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
 Tel: 33 1 34 65 25 16
 Fax: 33 1 34 65 25 21
 Email: sorokine@jouy.inra.fr
 best homologue in strain IL1403 is nadR (98%)
 Class: shotgun
 High quality sequence start: 30
 High quality sequence stop: 3207.
 FEATURES
 source
 1. .3237
 /organism="Lactococcus lactis subsp. cremoris"
 /mol_type="genomic DNA"
 /strain="MG1363"
 /sub_species="cremoris"
 /db_xref="taxon:1359"
 /clone_lib="MG1363 Random Sequence Tag Library"
 /notes="Vector: pSGM02; Site 1: Smal; Library of chromosomal fragments of L.lactis strain MG1363 was prepared by partial AluI digestion or by sonication."

ORIGIN

Query Match 3.1%; Score 73.2; DB 28; Length 3237;
 Best Local Similarity 49.8%; Pred. No. 9.5e-08;
 Matches 287; Conservative 0; Mismatches 263; Indels 26; Gaps 3;

1805 AACTCGGCGTCAGAAATCCAGGTTCTGTCAGTGAATCAAGGCTGTGTCCAGCTGCTCA 1864
 Db
 1956 AATTCTTCGTGAGTAATTCAGAGTTCTTAAGCTGCAATTTTCATGATGTTGTTCG 2015
 1865 ACTGACGAAGCACAATCAATGACACTGGTCAAGTATCGGCGCTACTCTCTGCTGC 1924
 Db
 2016 GGACGACTTGGCCCAATCAAGCTGATGAACAGCTTGAACCTTTGCCAGCTCTT---TCC 2072
 1925 CGGACGACCCATGACAGCCCATCTGCGAAGTACTGCGCGCTTCTGGCGGATGCA 1984
 Db
 2073 CGTAAACCCCAAGCTTATGCCAATTTGAGCTAGAGATTGTCCACGACTTTGAGCTAAATCA 2132
 1985 TTGAGCTTGGGACATATCAATATTTTCAGTTTCAATGTCATGCGCTCCCTCAGACAGGACTTA 2044
 Db
 2133 TTAAGGCTTGAATTTGTTCCAACTTTCTGTGTTAAGACTGTCAATCATGCAAGTCA 2192
 2045 CCTCGCTGGCGGG-----GAACCCCTCTGGAATTTCAATCGAGATATTGTCCGTGAGC 2098
 Db
 2193 TAATGAGGATCAGCAATTCGCAATTTTCAGGAATTTCCATGCAATATATTTTCTGTTAAG 2252
 2099 AGGCCCTGCGCAAGTGTGTGAGAACATGACCGCAAGACCATTTGTTGGCAGCTGACTGC 2158
 Db
 2253 AGTCTCTGATACAAAGGCTTTAAAGGCAATTTGTTCTTATCCC-----CC 2295
 2159 AACAGTTCTTCCAGCTTCATCGCCGCTTCTCCACCAACGATTAATGATGAATAGCTT 2218
 Db
 2296 CCTCTGTCAAAGTTCTTGTAAATCATCTTCAATCCATCGATCTAACATTGAATAAGCA 2355
 2219 GGCTGATGAATCAGAAGCGGGGAGCCCTCTCTCGCGCATGAATCAGCGGCTCCGCTGTG 2278
 Db
 2356 GGCTGATGAATCAAGAGTTTAAAGGCTTAACTTTTTCAGCGCGCTTAAGACCGCAGCTCGGTT 2415
 2279 AGCTCTGACCGTAGGAGGAATATCCACGTTAAGAGGCTTTCCAGAGCGCAACAATGTCA 2338


```

Db      2416 TCTTGGCAGAAATATAGACAAACCCATAAAGTCTTTTCCATATCAGAGCAGTT 2475
Qy      2339 CGCAATGCGTACATGGTTCTTCCAAAGGAGTATCT 2374
Db      2476 TTCAAGGCGCCATGGTTCTTCAATAGGGGTATTT 2511

RESULT 3
AF075981/c
LOCUS      AF075981
DEFINITION      AF075981 Salmonella typhimurium LT2, Lambda DASH II Salmonella
                typhimurium genomic clone 390-T3, genomic survey sequence.
ACCESSION      AF075981
VERSION        AF075981.1 GI:3320851
KEYWORDS       GSS.
SOURCE         Salmonella typhimurium
ORGANISM       Salmonella typhimurium
               Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
               Enterobacteriaceae; Salmonella.
REFERENCE      1 (bases 1 to 860)
AUTHORS        Wong R.M.-Y., Wong K.K., Benson N.R. and McClelland M.
TITLE          Sample sequencing of a Salmonella typhimurium LT2 lambda library:
               comparison to the Escherichia coli K12 genome
JOURNAL        FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
MEDLINE        99243757
PubMed        10227170
COMMENT        Contact: McClelland M
                Molecular Biology
                Sidney Kimmel Cancer Center
                3099 Science Park Road, San Diego, CA 92121, USA
                Email: mclelland@lifsci.sdsu.edu
                Class: shotgun.
FEATURES       Location/Qualifiers
                source
                1. .860
                /organism="Salmonella typhimurium"
                /mol_type="genomic DNA"
                /strain="LT2"
                /db_xref="taxon:602"
                /clone="390-T3"
                /clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
                /note="Vector: Lambda DASH II; sequenced using Li-Cor
                sequencer"

ORIGIN
Query Match      3.0%; Score 72; DB 28; Length 860;
Best Local Similarity 54.7%; Pred. No. 9.9e-08;
Matches 220; Conservative 0; Mismatches 160; Indels 22; Gaps 3;

Qy      1925 GCAGACCCATCGACGCGCATCTGCGCAGTGAAGTCCGCGGTTCTCGGCGCATGTCA 1984
Db      400  CGCAGATCCGCGGAGCGCCATTTGATCAATTTTGAACCCGTCGCGCAGCGTTC 341

Qy      1985 TTGAGCTTGGCAGCATATCAATATTGTTTCAAGTCCATCGCCCTCAGACAGGAGTTA 2044
Db      340  TTCAACCGGAGCAATTTTTCAGTTTCTCGCGGTAATCTGTTCTGTTTAAAGAAACGG 281

Qy      2045 CCCTGGGTG---GCGGCGGAACCTCTGGAATTCATCGAGATATTTGTCGCTGAGCAGG 2101
Db      280  CTTCCGCTCGCCGCGCGGGAATTTTCGGAATACCAATTCAAATTACGCTGATGAGC-TG 222

Qy      2102 CCCTGCGCAAGTGTGAGAAAGCAATGAGCCAGACCATTTGTTGGCAGTCACTGCAAC 2161
Db      221  CCGCGCCCGAGCGGAGAGAGCAATACTACGAGCTTTTTCCTGCGAAGCGCCAGC 162

Qy      2162 AAGTTCTTCAACGTCATCGCCCGGTTCTCTCCACCCCAAGATTAATGATGAAATAGCTTGGC 2221
Db      161  AG-----CCCGTCTCTTACCCAAAGCTCAAAAAGCGAATATTTAGGC 120

Qy      2222 TGATGATCAAGAGCGGAGCGCTCTCTCGCGCATGAAGTCAAGCGCGCTCCGCTGTGAGC 2281
Db      119  TGATGAATCAGGAGCGGTGCGGAGATCTCTCAGAAATATCGATAGCTTGTCTGCGCCAGA 60

```

```

Qy      2282 TCTGACCGTAGAAGAAATACCCAGCTAAAGAGCCTTTCCA 2323
Db      59  TCGCAGGATATTAGATCCCCACGATACACGCTTTGCCA 18

RESULT 4
CNS01MQH/c
LOCUS      CNS01MQH
DEFINITION      CNS01MQH 830 bp DNA linear GSS 14-JUN-2001
                Anopheles gambiae GSS T7 end of clone 22E24 of NotreDamel library
                from strain PEST of Anopheles gambiae (African malaria mosquito),
                genomic survey sequence.
ACCESSION      AL151258
VERSION        AL151258.1 GI:7011737
KEYWORDS       GSS.
SOURCE         Anopheles gambiae (African malaria mosquito)
ORGANISM       Anopheles gambiae
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
               Anopheles.
REFERENCE      1 (bases 1 to 830)
AUTHORS        Genoscope.
TITLE          Direct Submission
JOURNAL        Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : Segref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
REFERENCE      2 (bases 1 to 830)
AUTHORS        Roth C.W., Brey P.T., Ke Z., Collins F.H. and Weissenbach J.
TITLE          Direct Submission
JOURNAL        Submitted (16-FEB-2000) BEM, Institut Pasteur, 25, rue du Dr.
                Roux, Paris 75015, France
COMMENT        This clone is from an A. gambiae BAC library provided by F.H.
                Collins and sequenced by Genoscope in collaboration with the
                Laboratory of Biochem. and Biol. Molec. of Insects, Institut
                Pasteur.
FEATURES       Location/Qualifiers
                source
                1. .830
                /organism="Anopheles gambiae"
                /mol_type="genomic DNA"
                /strain="PEST"
                /db_xref="taxon:7165"
                /clone="22E24"
                /clone_lib="NotreDamel"
                /note="end : T7"

ORIGIN
Query Match      2.5%; Score 60.4; DB 29; Length 830;
Best Local Similarity 50.7%; Pred. No. 0.00013;
Matches 142; Conservative 1; Mismatches 137; Indels 0; Gaps 0;

Qy      1038 TTACAGGTCTGCTTTTGGGGCGCAGTCTTTTACTGTCATCGACCGCAGAAATGACTGG 1097
Db      597  KTCAGGGCTGCTTTAGTGCATTTTATTCTTCCCTCGCGCCCAAAATGGTTTCG 538

Qy      1098 TGATTAACAAGAAATTAAGCGGCAAGACTCATTCGCGGTTCTTCGTCGTTAAATT 1157
Db      537  TGATGAATCAGGGCATCGCCGCCAGTACCATCTGATGATTCGCTGCTATGCGGGTAA 478

Qy      1158 CTGACGCTTTTGTTCATCGCGCCACCTTGGGGGTTGATCTTTTGTCCAATGCGCGC 1217
Db      477  GTGATTTGCTGTAATCTGTGCGCGGATTTTGGCGGCGCGCTGCTGATGAGTCTC 418

Qy      1218 CGATCGTCTCGATATTATATGCGCTGGGGTGGCATCGCTTACCTTTATGTTTGGCGTCA 1277
Db      417  CGTGGCTGCTGGCGTTGGTGGTCCACTGGGGCGGCTGGCTTCTGCTCTGTCGATCGGATTCG 358

Qy      1278 TGGCAGCGAAGAGCGCCATGACAAACAGGTGGAAGCGCC 1317
Db      357  GCGCGCTGAAACGCGGATGACAGTAACTCGAACTGGC 318

RESULT 5
B2554294
LOCUS      B2554294
                1034 bp DNA linear GSS 17-DEC-2002

```

```

DEFINITION pacsl-60_4617.xl pacsl-60 Pseudomonas aeruginosa genomic clone
ACCESSION BZ554294
VERSION BZ554294.1 GI:27161466
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1034)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
1. .1034
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacsl-60_4617"
/clone_lib="pacsl-60"
/notes="clinical isolate 1-60 Whole genomic shotgun
library."
ORIGIN
Query Match 2.4%; Score 56.8; DB 28; Length 1034;
Best Local Similarity 50.4%; Pred. No. 0.0014;
Matches 139; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
QY 1022 ATCATGGAAATCTTCAATACAGTCTGCTTTGGGGCCAGTCTTTTACTGTCACATCGGA 1081
DB 82 ATGTTTCAGAGCTATCTCAAGCATCTCTGGTGGCTGGCGGCTGATATCGCCATTCGC 141
QY 1082 CCGCAGATGCTACTGTGATTAACAAGAAATTAAGCGGGAAGACTCAATTCGGTCTTT 1141
DB 142 GCGGAGATGCTTCTGCTTCGCGAGAGCTCGCGGAGCATCACTCTCGGTAGCC 201
QY 1142 CTCGTGTGTTTAAATTTCTGACGTCTTTTCTCATCGCCGGGACCTTGGGCGGTGATCTT 1201
DB 202 GCGCTCTGCGTGTCTCGACGCGGTGCTGCTCAGCTCGCGGTGCTTGGCCCAAG 261
QY 1202 TTGTCCATGCGCGGCGATCGTCTCGATATATATCGTGGGTGGCATCGTTACCTG 1261
DB 262 CTGCTGTGGAACACCGACGCTGCTGGCCATCGCCGCTGGGGCGAATCGGCTTCTG 321
QY 1262 TTATGTTTGGCGTCAATGAGCAGCAAGACGCCATG 1297
DB 322 ACCTGGTAGGGCTCAAGGGCTGCTTCGGCGGTG 357
RESULT 6
BZ560550/c
LOCUS BZ560550 897 bp DNA linear GSS 17-DEC-2002
DEFINITION pac82-164_2569.xl pac82-164 Pseudomonas aeruginosa genomic clone
pac82-164_2569, genomic survey sequence.
ACCESSION BZ560550
VERSION BZ560550.1 GI:27178749
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 897)

```

```

AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
1. .897
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac82-164_2569"
/clone_lib="pac82-164"
/notes="clinical isolate 2-164 Whole genomic shotgun
library."
ORIGIN
Query Match 2.3%; Score 54.2; DB 28; Length 897;
Best Local Similarity 52.8%; Pred. No. 0.0067;
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;
QY 1455 TGATGGCAATCGTCTGACCTGGTTGAACCGAATTCGCTATTTCGACGCGTTTGTGTTTA 1514
DB 552 TCGGGGCACTGGGCTTCACTGGCTTAACCCCACTCTATTTCATACGCTACTTCTTAC 493
QY 1515 TCGGGGCGTTCGGGCAATACGCGCACACCGGACGCTGGAATTTTCGCGTGGCGGT 1574
DB 492 GGTTCGCTCGCGCCAGCAAGNCGCG-----CCGGGCGCTATGCTCGCGCGG 441
QY 1575 TCGCGCAAGCTGATCTGCTTCGCTGCTGGTTCGGCGCAGCAGCATTTGTCAACGC 1634
DB 440 CCAAGCGCTCGCTGATGTGTTCTTCGCGCTCGCCCTCGCGCGGCGCATGGCTGGCCCT 381
QY 1635 CGCTGTCAGCCCCCAAGGTGCGCTGATCAACCTCGTGTGCGAGTTGTGATGACCG 1694
DB 380 GGTTCGCGCGCGCGGCACTCGCGCCCTGCTCGACTGATGATGCGCGCCATGATGCTGG 321
QY 1695 CATTCGCCATCAAACTGATGTTGATGGTTA 1725
DB 320 GCATGCGCGCGCAACTGCTGTTCCGGGATA 290
RESULT 7
BZ568946/c
LOCUS BZ568946 1620 bp DNA linear GSS 17-DEC-2002
DEFINITION pac82-164_8165.y2 pac82-164 Pseudomonas aeruginosa genomic clone
pac82-164_8165, genomic survey sequence.
ACCESSION BZ568946
VERSION BZ568946.1 GI:27202770
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1620)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954

```

```
aps 0;
```

ACCT 1671

9

```

ORIGIN
Query Match          2.0%; Score 48.6; DB 28; Length 833;
Best Local Similarity 50.8%; Pred.No. 0.21;
Matches 117; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
/clone.lib="pacsl-60"
/notes="clinical isolate 1-60 Whole genomic shotgun
library."

```

QY 1067 TTAATGTCATCGGACCGGAGATGATCTGTTGATTAACAGAGATTAAGCGGAGCA 1126
 Db 446 TCATCGGACCATCGGCGGAGAGATGCTTCTGCTCGGAGAGCGCTCGCGCGAGCAT 387
 QY 1127 CTCATTCGGGTTCTTCTGCTGCTTTAAATCTGACCTCTTTTGTTCATCGCGGCA 1186
 Db 386 CACCTCTCGGTAGCGGCTCTGCTGTTCTGCGACACGCTGCTGCTGAGCTCGGCTG 327
 QY 1187 TTGGGCGTGTGATCTTTTCTCAATGCGCGCGCATGCTGCTGATATATCGCTGGGTT 1246
 Db 326 TTGGGCTGCGCAAGCTCTGCTGGAACCGGAGCTGCTGCGCATCGCGCTGGGCG 267
 QY 1247 GCGATCGCTTACTGTTATGTTTTCGCTGATGACGAGCAAGAGCCCATG 1297
 Db 266 GGGATCGCTTCTGACCTGATGAGGCTCAAGGCGCTGCTTTCGCGGTTG 216

RESULT 10
 CNS0091P 925 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
 DEFINITION BAC191916 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 ACCESSION AL053013 GI:4934461
 VERSION AL053013.1
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 925)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr)
 COMMENT - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org/TheBDGP/Drosophila>
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 source
 1..925
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BAC191916"
 /clone_lib="RPCI-98"
 /note="end : TET3"

ORIGIN
 Query Match 2.0%; Score 48.6; DB 29; Length 925;
 Best Local Similarity 13.8%; Pred. No. 0.22; Indels 0; Gaps 0;
 Matches 49; Conservative 157; Mismatches 149;
 QY 1340 CCAACGCTGCCGATGACAGCCTTTTGGCGGTTTCGCGGCTGACACGCGCAAC 1399
 Db 571 SCSCGSCGSCSSSCSCBCCCCSSSVCCSSSSSKSSTSSSCSSKSVCGTS 630
 QY 1400 CGGTCGCGGTGAGGTGACGTCGATAGACGCGGTTTGGGTAAACCCATGTTGATG 1459
 Db 631 CSSSSSCSSSSSSTSSSTSSSTSSSSSGSSSSSSTTSKTSASGSGSWSAGGSG 690

QY 1460 GCAATCGTCTGACCTGTTGAACCCGGAATGCTATTGGACGCGTTTGTGTTATCGGC 1519
 Db 691 BTGTSSTSSSSSTSTSSSVSGSKSTSSGSSSSSSSTSSBBSCTSTSSSS 750
 QY 1520 GCGCTCGGCGCGCAATACGGGACACCGGACCGGTGATTTTCCGCTCGCGCTGCGG 1579
 Db 751 SSVSSSTCSCCTCCCSYSSVSSSTSSSSSTSGSSSSSVGTSSSSSDSTSTCCSCCY 810
 QY 1580 GGAAGCTGATCTGTTTCCGCTGTTGCGGTCGCGCAGCAGCATTGTCAAGCCGCTG 1639
 Db 811 MCTCCSTYBCTSTSSGSSSSGKGVTGCGGCGGSSSTNMBGTSSNACSSSSSC 870
 QY 1640 TCCAGCCCCAAGGTGCGCTGATCAACGCTGCTGCTGCGCAGTGTGTGATGACCG 1694
 Db 871 SSSSVSSSSKSSASSSSVSSGSSSVSSSSASKSSSSGSSSVSSGSGSGSVSG 925

RESULT 11
 BX393687 852 bp mRNA linear EST 13-MAY-2003
 LOCUS BX393687 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DC001YG07 5-PRIME, mRNA sequence.
 ACCESSION BX393687
 VERSION BX393687.1 GI:30624032
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 852)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Paraday Avenue Genoscope, sequence ID : CS0DC001AD04QF1.

FEATURES
 source
 1..852
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC001YG07"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 2.0%; Score 47.6; DB 13; Length 852;
 Best Local Similarity 17.2%; Pred. No. 0.4;
 Matches 69; Conservative 152; Mismatches 178; Indels 3; Gaps 1;
 QY 1038 TTACAGTCTGTTTGGGGCGCGCTTTTACTGCGACCGCAGAGTACTCGG 1097
 Db 381 TTTCKGKTBKTKTKTKGKKDKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 440
 QY 1098 TGATTAACACAGGAATTAAGCGGAGGACTCATTCGCGGTCTTCTGCGTGTGTTAATT 1157
 Db 441 KKK 500
 QY 1158 CTGACGCTCTTTTGTTCATCGCGGACCTTGGGGGCTGTGATCTTTTGTCCAGTCCGCGC 1217
 Db 501 KDKKK 560
 QY 1218 CGATCGTCTCGATATTATTCGCGCTGGGCTGGCATCGCTTACCTGTTATGTTTCCGCTCA 1277

Thu Mar 18 12:31:00 2004

LOCUS
DEFINITION BQ704105 644 bp mRNA linear EST 16-JUL-2002
346110C04.y1 946 - tassell primordium prepared by Schmidt lab Zea
mays cDNA, mRNA sequence.

ACCESSION BQ704105
VERSION BQ704105.1 GI:21843524
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
AUTHORS Walbot.V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946110 row: C column: 04.
Location/Qualifiers
1. 644
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="X10LR"
/clone_lib="946 - tassell primordium prepared by Schmidt
lab"

FEATURES
source

ORIGIN

Query Match 1.9%; Score 44.4; DB 13; Length 644;
Best Local Similarity 46.0%; Pred. No. 2.5;
Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

Qy 1364 TTGGCGGTTTCGGCGGTCGACACCGGCAACCGGTCGGGTCGAGTGACGCTC 1423
Db 364 TTGGCGGTTTCGGCGGTCGACACCGGCAACCGGTCGGGTCGAGTGACGCTC 305

Qy 1424 GATAAGCAGCGGTTTGGGTAAGCCCATGTTGATGGCAATCGTGTGACCTGTTGAAC 1483
Db 304 CTCCTCTAGTAGACGAGTGGGAGCCGAAGTGGACGTAGCAGGTGCGCGGAGTGATC 245

Qy 1484 CCGAATGCGTATTTCGACGCGCTTTGTTTATTCGCGGCGCTCGCGCGCATACGCGAC 1543
Db 244 TCGAAGCGCGGTCGCTGCTGGAAGCTGTAGGAGCGGAGCGGAGTTCGCGGATGAGCCCC 185

Qy 1544 ACCGACGCGTGGATTTTCGCGGTCGGGCTTCGCGGCAAGCTGATCTGCTCCGCTG 1603
Db 184 TTCGGGAGGCGGCTTCTGGGAGTAGTCTTTCGCGGCGGCTGAGCGAGCTGCTGAC 125

Qy 1604 GTGGGTTTCGGCGGACGACGATTGTTCACGCGCTGTTCAGCCCCCAAGGTGTGGCGTGG 1663
Db 124 GCCGTGACCGCTCGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 65

Qy 1664 ATCAACGCTCGTGTGGCAGTTGTGAT 1689
Db 64 GACATTGCGGATATTGGGAT 39

Search completed: March 16, 2004, 02:52:12

LOCUS
DEFINITION BQ704105 644 bp mRNA linear EST 16-JUL-2002
346110C04.y1 946 - tassell primordium prepared by Schmidt lab Zea
mays cDNA, mRNA sequence.

ACCESSION BQ704105
VERSION BQ704105.1 GI:21843524
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
AUTHORS Walbot.V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1114036 row: D column: 07.
Location/Qualifiers
1. 620
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone_lib="1114 - Unigene IV from Maize Genome Project"
/note="This library represents the unique genes found in
the fourth round of EST sequencing at Stanford University
for the maize genome project. Sequences are present from
libraries 1091 and 3524. Contigs were assembled using
ZnDBAssembler and 2 representatives from each contig were
selected for the Unigene set. All singlets were also
selected."

FEATURES
source

ORIGIN

Query Match 1.9%; Score 44.4; DB 14; Length 620;
Best Local Similarity 46.0%; Pred. No. 2.5;
Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

Qy 1364 TTGGCGGTTTCGGCGGTCGACACCGGCAACCGGTCGGGTCGAGTGACGCTC 1423
Db 357 TTGGCGGTTTCGGCGGTCGACACCGGCAACCGGTCGGGTCGAGTGACGCTC 298

Qy 1424 GATAAGCAGCGGTTTGGGTAAGCCCATGTTGATGGCAATCGTGTGACCTGTTGAAC 1483
Db 297 CTCCTCTAGTAGACGAGTGGGAGCCGAAGTGGACGTAGCAGGTGCGCGGAGTGATC 238

Qy 1484 CCGAATGCGTATTTCGACGCGCTTTGTTTATTCGCGGCGCTCGCGCGCATACGCGAC 1543
Db 237 TCGAAGCGCGGTCGCTGCTGGAAGCTGTAGGAGCGGAGGTCGCGGATGAGCCCC 178

Qy 1544 ACCGACGCGTGGATTTTCGCGGTCGGGCTTCGCGGCAAGCTGATCTGCTCCGCTG 1603
Db 177 TTCGGGAGGCGGCTTCTGGGAGTAGTCTTTCGCGGCGGCTGAGCGAGCTGCTGAC 118

Qy 1604 GTGGGTTTCGGCGGACGACGATTGTTCACGCGCTGTTCAGCCCCCAAGGTGTGGCGTGG 1663
Db 117 GCCGTGACCGCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 58

Qy 1664 ATCAACGCTCGTGTGGCAGTTGTGAT 1689
Db 57 GACATTGCGGATATTGGGAT 32

Search completed: March 16, 2004, 02:52:12

RESULT 15
BQ704105/c

Job time : 6533.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:15:20 ; Search time 2846.7 Seconds
(without alignments)
10825.501 Million cell updates/sec

Title: US-09-105-117K-1 COPY 1016 1726

```

acc: 33 33 105 117K_10611_1120
perfect score: 711
Sequence: 1 atgtdaatcatggaatctt.....aactgatcttcatgggttag 711

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_gr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_or:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
C 1	711	100.0	711	6	AX643028	AX643028 Sequence	
	2	711	100.0	822	AX063771	AX063771 Sequence	
	3	711	100.0	822	AX244059	AX244059 Sequence	
	4	711	100.0	2374	1	CGLYS8G	X96471_C.Glutamicu
	5	711	100.0	2374	6	AX93933	AX93933 Sequence 2
C 6	711	100.0	333150	1	AP005277	AP005277 Corynebac	
	7	711	100.0	349980	6	AX127147	AX127147 Sequence
	8	708	99.6	708	6	AX123539	AX123539 Sequence
	9	708	99.6	708	6	BD165656	BD165656 Novel pol
	10	696.8	98.0	712	6	AX643030	AX643030 Sequence
C 11	317.6	44.7	1568	6	E54483	E54483 Heat-resist	
	12	317.6	44.7	1771	1	AX083133	AX083133 Corynebac
	13	317.6	44.7	308650	1	AP005218	AP005218 Corynebac
	14	136.6	19.2	349535	1	BX248357	AX248357 Corynebac
	15	84.8	11.8	190050	1	AX1646059	AX1646059 Ralstonia
C 16	65.6	9.3	300861	1	AE016777	AE016777 Pseudomon	
	17	65.6	9.2	303550	1	SC0939131	AL939131 Streptomy
	18	63.4	8.9	10542	1	AE004852	AE004852 Pseudomon
	19	63.2	8.9	10024	1	AE015725	AE015725 Shewanell
	20	62.8	8.8	24336	1	AE008841	AE008841 Salmonell
C 21	62.8	8.8	298300	1	AP005025	AP005025 Streptomy	
	22	62.2	8.7	311600	1	AE016871	AE016871 Pseudomon
	23	62	8.7	10313	1	AE015304	AE015304 Shigella
	24	62	8.7	10701	1	AE005522	AE005522 Escherich
	25	62	8.7	266658	1	AP002563	AP002563 Escherich
C 26	62	8.7	292906	1	AE016988	AE016988 Shigella	
	27	61.6	8.7	4033	1	ASU65741	U65741 Aeromonas s
	28	61.2	8.6	230050	1	AL672727	AL672727 Salmonell
	29	61.2	8.6	301574	1	AE016844	AE016844 Salmonell
	30	60.4	8.5	636	6	E49392	E49392 Process for
C 32	60.4	8.5	636	6	AX030085	AX030085 Sequence	
	33	60.4	8.5	8029	1	ECFDAP3K	X14436 Escherichia
	34	60.4	8.5	10362	1	AE000375	AE000375 Escherich
	35	60.4	8.5	141744	1	ECU28377	U28377 Escherichia
	36	60	8.4	211182	2	AL646067	AL646067 Ralstonia
C 37	59.6	8.4	10747	1	AE012410	AE012410 Xanthomon	
	38	58.8	8.3	303121	1	AE016766	AE016766 Escherich
	39	57	8.0	14844	1	MBU34849	U34849 Mycobacteri
	40	57	8.0	14869	1	AE007056	AE007056 Mycobacte
	41	57	8.0	15239	6	186263	186263 Sequence 17
C 42	57	8.0	306050	1	EX248341	EX248341 Mycobacte	
	43	57	8.0	346186	1	BX842578	BX842578 Mycobacte
	44	56.6	8.0	642	6	AX387945	AX387945 Sequence
	45	56.6	8.0	248650	1	AP005341	AP005341 Vibrio vuln

ALIGNMENTS

RESULT 1					
AX643028					
LOCUS		711 bp	DNA	linear	PAT 24-FEB-2003
DEFINITION	Sequence 7 from Patent EP1266956.				
ACCESSION	AX643028				
VERSION	AX643028.1	GI:28550158			
KEYWORDS					
SOURCE	Corynebacterium glutamicum				
ORGANISM	Corynebacterium glutamicum				
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
	Corynebacterineae; Corynebacteriaceae; Corynebacterium.				
REFERENCE 1					
AUTHORS	Gunji, Y. and Yasueda, H.				
TITLE	Method for producing L-lysine or L-arginine by using methanol assimilating bacterium				

JOURNAL Patent: EP 1266966-A 7 18-DEC-2002;
 Ajinomoto Co., Inc. (JP)
 FEATURES Location/Qualifiers
 source 1..711
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 CDS 1..711
 /notes="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD67811.1"
 /db_xref="GI:28550159"
 /translation="MEIFITGLGASLLSISGPNVIVIKQIKRGLIIVLVCL
 CLISDVFLFAGTIGVLDLSNAAPVILDMRWGGIAYLLWFAVMAKADMTNKVEAPQ
 IIEETFTVPDTPGGSVATDTRNRVREVSVDKORVWVKPMLMAIVLTWLPNAYLDA
 LDAFVIGVGAQVGDTRWIFAAAGFAASLIWPLVFGFAAALSRLSPSSPKWRINVV
 VVAVVMTALAIKLMNG"
 ORIGIN
 Query Match 100.0%; Score 711; DB 6; Length 711;
 Best Local Similarity 100.0%; Pred. No. 2.6e-169;
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGTGATCATGAATCTTCAATACAGGTCGCTTTGGGGCCAGTCTTTTACTGTCC 60
 DB 1 ATGGTGATCATGAATCTTCAATACAGGTCGCTTTGGGGCCAGTCTTTTACTGTCC 60
 QY 61 ATCGGACCGCAGAAATCTACTGTGATTAAACAAGGAATTAAGCGCGAAGGACTCATTTGCG 120
 DB 61 ATCGGACCGCAGAAATCTACTGTGATTAAACAAGGAATTAAGCGCGAAGGACTCATTTGCG 120
 QY 121 GTTCTTCTGCTGTGTTAAATTTCTGACGTCCTTTTGTTCATCGCCGCGACCTTGGCGGTT 180
 DB 121 GTTCTTCTGCTGTGTTAAATTTCTGACGTCCTTTTGTTCATCGCCGCGACCTTGGCGGTT 180
 QY 181 GATCTTTTGTCCAAATCGCGCGCATCGTCTCGATATTATGCGCTGGGTCGCGATCGCT 240
 DB 181 GATCTTTTGTCCAAATCGCGCGCATCGTCTCGATATTATGCGCTGGGTCGCGATCGCT 240
 QY 241 TACCTGTGTATGGTTTTCGCTCATGCGCAGGAAAGCGCATGACAAACAAGTGTGAAGCG 300
 DB 241 TACCTGTGTATGGTTTTCGCTCATGCGCAGGAAAGCGCATGACAAACAAGTGTGAAGCG 300
 QY 301 CCACAGATCATTAAGAAACAGAACCAACCGTCCCGGATGACGCGCTTTGGCGGTTTCG 360
 DB 301 CCACAGATCATTAAGAAACAGAACCAACCGTCCCGGATGACGCGCTTTGGCGGTTTCG 360
 QY 361 GCGGTGCGCACTGACACGCGCAACCGGTCGCGGTGGAGTGGAGTGGATGAAGCAGCGG 420
 DB 361 GCGGTGCGCACTGACACGCGCAACCGGTCGCGGTGGAGTGGAGTGGATGAAGCAGCGG 420
 QY 421 GTTTGGGTAAAGCCCAATGTTGATGCGCAATCGTGTGACCTGGTGGGTTCGCG 480
 DB 421 GTTTGGGTAAAGCCCAATGTTGATGCGCAATCGTGTGACCTGGTGGGTTCGCG 480
 QY 481 TTGGACGGTTTGTGTTTATCGCGCGTCCGCGCAATACGCGGACACCGGACGTTGG 540
 DB 481 TTGGACGGTTTGTGTTTATCGCGCGTCCGCGCAATACGCGGACACCGGACGTTGG 540
 QY 541 ATTTTTCGCGCTTGGCGGTTTCGCGCAAGCTGATCTGGTTCGCGCTGGTGGGTTCGCG 600
 DB 541 ATTTTTCGCGCTTGGCGGTTTCGCGCAAGCTGATCTGGTTCGCGCTGGTGGGTTCGCG 600
 QY 601 GCAGCAGATGTTGACGCGCGCTGTTCAGCCCCCAAGGTGTGCGCTGGATCAACGTCGTC 660
 DB 601 GCAGCAGATGTTGACGCGCGCTGTTCAGCCCCCAAGGTGTGCGCTGGATCAACGTCGTC 660
 QY 661 GTGCGAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 711
 DB 661 GTGCGAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 711

RESULT 2
 LOCUS AX063771 822 bp DNA linear PAT 24-JAN-2001
 DEFINITION Sequence 53 from Patent WO0100843.
 ACCESSION AX063771
 VERSION AX063771.1 GI:12541483
 KEYWORDS Corynebacterium glutamicum
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1
 AUTHORS Pompejus M., Kroeger B., Schroeder H., Zelder O. and Haberhauer G.
 TITLE corynebacterium glutamicum genes encoding metabolic pathway
 proteins
 JOURNAL Patent: WO 0100843-A 53 04-JAN-2001;
 BASF AKTIENGESELLSCHAFT (DE)
 FEATURES Location/Qualifiers
 source 1..822
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 CDS 101..802
 /notes="unnamed protein product; RXA01394"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC25127.1"
 /db_xref="GI:12541484"
 /translation="MEIFITGLGASLLSISGPNVIVIKQIKRGLIIVLVCL
 SDVFLFAGTIGVLDLSNAAPVILDMRWGGIAYLLWFAVMAKADMTNKVEAPQ
 IIEETFTVPDTPGGSVATDTRNRVREVSVDKORVWVKPMLMAIVLTWLPNAYLDA
 LDAFVIGVGAQVGDTRWIFAAAGFAASLIWPLVFGFAAALSRLSPSSPKWRINVV
 AVVMTALAIKLMNG"
 ORIGIN
 Query Match 100.0%; Score 711; DB 6; Length 822;
 Best Local Similarity 100.0%; Pred. No. 2.7e-169;
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGTGATCATGAATCTTCAATACAGGTCGCTTTGGGGCCAGTCTTTTACTGTCC 60
 DB 92 ATGGTGATCATGAATCTTCAATACAGGTCGCTTTGGGGCCAGTCTTTTACTGTCC 151
 QY 61 ATCGGACCGCAGAAATCTACTGTGATTAAACAAGGAATTAAGCGCGAAGGACTCATTTGCG 120
 DB 152 ATCGGACCGCAGAAATCTACTGTGATTAAACAAGGAATTAAGCGCGAAGGACTCATTTGCG 211
 QY 121 GTTCTTCTGCTGTGTTAAATTTCTGACGTCCTTTTGTTCATCGCCGCGACCTTGGCGGTT 180
 DB 212 GTTCTTCTGCTGTGTTAAATTTCTGACGTCCTTTTGTTCATCGCCGCGACCTTGGCGGTT 271
 QY 181 GATCTTTTGTCCAAATCGCGCGCATCGTCTCGATATTATGCGCTGGGTCGCGATCGCT 240
 DB 272 GATCTTTTGTCCAAATCGCGCGCATCGTCTCGATATTATGCGCTGGGTCGCGATCGCT 331
 QY 241 TACCTGTGTATGGTTTTCGCTCATGCGCAGGAAAGCGCATGACAAACAAGTGTGAAGCG 300
 DB 332 TACCTGTGTATGGTTTTCGCTCATGCGCAGGAAAGCGCATGACAAACAAGTGTGAAGCG 391
 QY 301 CCACAGATCATTAAGAAACAGAACCAACCGTCCCGGATGACGCGCTGGATGAAGCAGCGG 360
 DB 392 CCACAGATCATTAAGAAACAGAACCAACCGTCCCGGATGACGCGCTGGATGAAGCAGCGG 451
 QY 361 GCGGTGCGCACTGACACGCGCAACCGGTCGCGGTGGAGTGGAGTGGATGAAGCAGCGG 420
 DB 452 GCGGTGCGCACTGACACGCGCAACCGGTCGCGGTGGAGTGGAGTGGATGAAGCAGCGG 511
 QY 421 GTTTGGGTAAAGCCCAATGTTGATGCGCAATCGTGTGACCTGGTGGGTTCGCG 480
 DB 512 GTTTGGGTAAAGCCCAATGTTGATGCGCAATCGTGTGACCTGGTGGGTTCGCG 571

```

QY 481 TTGGACGCGTTGTGTTTATCGCGCGCGTCCGCGCAATACGGCGACACCGACGCGTGG 540
Db 572 TTGGACGCGTTGTGTTTATCGCGCGCGTCCGCGCAATACGGCGACACCGACGCGTGG 631
QY 541 ATTTTCGCGCGTGGCGGCTTCGCGCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGCG 600
Db 632 ATTTTCGCGCGTGGCGGCTTCGCGCAAGCTGATCTGGTTCCCGCTGGTGGGTTTCGCG 691
QY 601 GCAGCAGCATTTGATGACCGCGCTGTCAGCGCCCAAGGTGGCGCTGGATCAACGTCGTC 660
Db 692 GCAGCAGCATTTGATGACCGCGCTGTCAGCGCCCAAGGTGGCGCTGGATCAACGTCGTC 751
QY 661 GTGCGAGTTGTGATGACCGCATTTGCGCATCAAACTGATGTTGATGGGTTAG 711
Db 752 GTGCGAGTTGTGATGACCGCATTTGCGCATCAAACTGATGTTGATGGGTTAG 802

RESULT 3
AX244059
LOCUS 822 bp DNA linear PAT 28-SEP-2001
DEFINITION Sequence 51 from Patent WO0166573.
ACCESSION AX244059
VERSION AX244059.1 GI:15859123
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS Pompeius, M., Kroeger, B., Schroeder, H., Zelder, O., Haberhauer, G.,
Kim, J.W., Lee, H.S. and Hwang, B.J.
TITLE Corynebacterium glutamicum genes encoding metabolic pathway
JOURNAL Patent: WO 0166573-A 51 13-SEP-2001;
BASIS AKTIENGESCHSCHAFT (DE)
FEATURES
source
1..822
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101..802
/note="unnamed protein product; RXA01394"
/codon_start=1
/transl_table=11
/protein_id="CAC88508.1"
/db_xref="GI:15859124"
/db_xref="REFSEQ:CAC88508"
/translations="MEIFITGLLLGASLLSTGPQNVLVIKQIKRGLIIVLVLCI
SDVFLIAGTLGVDLLSNAPIVLDIMRWGGIAYLLWFAVNAKADMTNKPAPQIE
ETEPTVDDPLGGSAAVATDRNRVSVQKRVKPKMLAIVLWLNPAFLDA
FVFIGGAGYGTGRWIPFAAGAFASLIWFPLVFGAAALSRLSPKVMRWINVVV
AVVMTALAIKMLMG"
CDS
ORIGIN
Query Match 100.0%; Score 711; DB 6; Length 822;
Best Local Similarity 100.0%; Pred. No. 2.7e-169;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGTGATCATCGAATCTTATTACAGGTCTGCTTTGGGGCGCAGCTTTTACTGTCC 60
Db 92 ATGGTGATCATCGAATCTTATTACAGGTCTGCTTTGGGGCGCAGCTTTTACTGTCC 151
QY 61 ATCGACCGCAGATGTACTGGTATTAAACAAGAAATTAAGCGCGAAGACTCATTTGG 120
Db 152 ATCGACCGCAGATGTACTGGTATTAAACAAGAAATTAAGCGCGAAGACTCATTTGG 211
QY 121 GTTCTCTCGTGTGTTAAATTTCTGACAGCTTTTGTTCATCGCGGACCTTGGGGGTT 180
Db 212 GTTCTCTCGTGTGTTAAATTTCTGACAGCTTTTGTTCATCGCGGACCTTGGGGGTT 271
QY 181 GATCTTTTCTCAATGCGCGCGCATCGTCTCGATATTATCGCTGGGTGGCATCGCT 240
Db 272 GATCTTTTCTCAATGCGCGCGCATCGTCTCGATATTATCGCTGGGTGGCATCGCT 331

```

```

QY 241 TACCTGTTATGTTTGGTTCATCGCGAGCAAGACGCGCATGACAAACAGGTGGAAGCG 300
Db 332 TACCTGTTATGTTTGGTTCATCGCGAGCAAGACGCGCATGACAAACAGGTGGAAGCG 391
QY 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGGGCGTTCG 360
Db 392 CCACAGATCATTTGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGGGCGTTCG 451
QY 361 CGGTTGGCACTGACACCGCGCAACCGGTTGGGTTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 420
Db 452 CGGTTGGCACTGACACCGCGCAACCGGTTGGGTTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 511
QY 421 GTTTGGGTTAAAGCCCATGTTGATGGCAATCTGTTGACCTGGTTGAACCCGATGCGGTAT 480
Db 512 GTTTGGGTTAAAGCCCATGTTGATGGCAATCTGTTGACCTGGTTGAACCCGATGCGGTAT 571
QY 481 TTGACGCGGTTTGTGTTATCGCGCGCGTCCGCGCAATACCGCGCAACCGCGACGTTGG 540
Db 572 TTGACGCGGTTTGTGTTATCGCGCGCGTCCGCGCAATACCGCGCAACCGCGACGTTGG 631
QY 541 ATTTTCGCGCGTGGCGGCTTCGCGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGCG 600
Db 632 ATTTTCGCGCGTGGCGGCTTCGCGCAAGCTGATCTGTTCCCGCTGGTGGGTTTCGCG 691
QY 601 GCAGCAGCATTTGTCAGCGCGCGTGTCCAGCCCAAGGTGGCGCTGGATCAACGTCGTC 660
Db 692 GCAGCAGCATTTGTCAGCGCGCGTGTCCAGCCCAAGGTGGCGCTGGATCAACGTCGTC 751
QY 661 GTGCGAGTTGTGATGACCGCATTTGCGCATCAAACTGATGTTGATGGGTTAG 711
Db 752 GTGCGAGTTGTGATGACCGCATTTGCGCATCAAACTGATGTTGATGGGTTAG 802

RESULT 4
CGLYSE
LOCUS 2374 bp DNA linear BCT 19-MAR-2001
DEFINITION C-glutamicum lyse and lyse genes.
ACCESSION X96471
VERSION X96471.1 GI:1729753
KEYWORDS lyse gene; lyse gene; Lysine export regulator protein; Lysine
exporter protein; lysine governor.
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS Vrljic, M., Sahm, H. and Eggeling, L.
TITLE A new type of transporter with a new type of cellular function:
L-lysine export from Corynebacterium glutamicum
JOURNAL Mol. Microbiol. 22 (5), 815-826 (1996)
MEDLINE 97126810
PUBMED 8971704
REFERENCE
AUTHORS Vrljic, M.M.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-1996) M.M. Vrljic, Institut fuer Biotechnologie
1, Forschungszentrum Juelich, Postfach 1913, D-52425 Juelich, FRG
FEATURES
source
1..2374
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/strain="R127"
/db_xref="taxon:1718"
complement(82..954)
/gene="lyse"
/gene="lyse"
/note="Lysine governor"
/codon_start=1
/transl_table=11
/product="Lysine export regulator protein"
/protein_id="CAA65323.1"

```



```

Db      1376  GCCTGGCCACTGACACGCGCAACCGGCTGCGGCTGAGGCTGAGCTGATGACGAGCGG 1435
QY      421  GTTGGGTAAGCCCATGTTGATGCGCAATCGTCTGACCTGCTGACCCGCAATCGGTAT 480
Db      1436  GTTGGGTAAGCCCATGTTGATGCGCAATCGTCTGACCTGCTGACCCGCAATCGGTAT 1495
QY      481  TTGGACGCGTTTGTGTTTATCGCGCGCGCTCGCGCGCAATACGCGCAACCGCAGCGGTGG 540
Db      1496  TTGGACGCGTTTGTGTTTATCGCGCGCGCTCGCGCGCAATACGCGCAACCGCAGCGGTGG 1555
QY      541  ATTTTCGCGCTGCGCGGCTTTCGCGGCAAGCTGATCGGTTCGCGCTGCGGTTCGCGC 600
Db      1556  ATTTTCGCGCTGCGCGGCTTTCGCGGCAAGCTGATCGGTTCGCGCTGCGGTTCGCGC 1615
QY      601  GCAGCAGCATTTGTCACCGCGCTGTCACGCGCCCAAGCTGCGCGCTGCGGTGATCAACGTCGC 660
Db      1616  GCAGCAGCATTTGTCACCGCGCTGTCACGCGCCCAAGCTGCGCGCTGCGGTGATCAACGTCGC 1675
QY      661  GTGGCAGTTGTGATGACCGCATTCGCGCATCAAACTGATGTTGATGGGTTAG 711
Db      1676  GTGGCAGTTGTGATGACCGCATTCGCGCATCAAACTGATGTTGATGGGTTAG 1726

RESULT 6
AP005277/c
LOCUS
DEFINITION Corynebacterium glutamicum ATCC 13032 DNA linear BCT 08-AUG-2002
4/10
ACCESSION AP005277 BA000036
VERSION AP005277.1 GI:21323710
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum ATCC 13032
Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
1
AUTHORS Nakagawa, S.
TITLE Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 333150)
AUTHORS Nakagawa, S.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.
Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,
Tokyo 194-8533, Japan (E-mail: snakagawa@xanagen.com,
Tel: 81-44-829-3031, Fax: 81-44-813-1651)
This sequence is conducted by collaboration of Kyowa Hakko Kogyo
Co. Ltd. And Kitasato University.
FEATURES
Location/Qualifiers
1..333150
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:196627"
/note="ATCC 13032"
complement(63..1292)
/gene="Cg10944"
complement(63..1292)
/gene="Cg10944"
/codon_start=1
/translation="Hypothetical membrane protein"
/protein_id="BAB98337.1"
/db_xref="GI:21323711"
/translation="MTNPTERNARRLLIHWGLQNIQDQIVAAKTVLPWLLQAAGP
FLIALVPIREAGMLPQRAITGVWVLRQTSKRVWVIGSNQFVSALGIVAAFLRGL
WALGITVLIALSLFRSMCSIASKDVQKVIKSKRGLVTRATVIGVWGLVAGL
ATAIFLGSHPFTRVLAIVASSFSMLFASIVFARLEYAKPATPKNAPSNPWRCII
AALKDKAFRRFVLVRSWMLVTALSTAFIVALAASGNSIDSLGFFLIASGLASWVG
RISGTWSDHSKNVMAGALFGSIVLILVLSAFAPAQINTLVFPLSPFLITLAHTA
ILVARKTVMDEAGQDRTYIVADANTIMGVLLIVGLSGFIAIFGNEALLFLAII
GLLTGISARGLKEVSAG"

gene
CDS
```

```

gene
CDS
1382..2917
/gene="Cg10945"
1382..2917
/gene="Cg10945"
/codon_start=1
/transl_table=11
/product="Putative multicopper oxidases"
/protein_id="BAB98338.1"
/db_xref="GI:21323712"
/translation="MPEINRRTFFKGAGVLAATVVCAGVLVACSSDDVRVCGEPRTL
PIPADLGRGSSVHFALQGESQILPDVTTKTFNGTHLGPLVVKGGDDVHV
DVNNLDMTVMWHGKMLPALADGSPHSPIGQPMSTMTVANDAAITLWHPHTHG
LTCGLAYEGLAGMIIVDEEATDKLDPREYGVDDIPLVMDHREFLEDGSLDEEDLPDL
BLQDTPFANGITNAHFADATTRVRFRVLNGSNMRYLAFSDRTTQVITASDGLLD
EGQDRTTLAIGPGEWEIIVLECEVDVTLGVSREDNVGVDDFVDPDFGMSDSFQL
LITGPSDDAAQAPALPGVLAKSTPEPIDATEFTFMTFSINDLQMDQBVDDVATL
HQQEVWIVTNDSDWPHNFHVDARVILKFGEDTVELFNDGKWDVTLGLPGATATL
AVEFGHPDPQPMYTHCMVLYHEDQGMQGFVIVEPGDPAALVIGSGSSIDSAGG
HAH"
3063..3983
/gene="Cg10946"
3063..3983
/gene="Cg10946"
/note="PF00005:ABC transporter
TIGR00960:3a0501s02: Type II (General) Secretory Pathway
(TISP) Family protein
TIGR00968:3a0106s01: sulfate transport system permease
protein
TIGR01166:cbiO: cobalt transport protein ATP-binding
subunit
TIGR01184:ntrCD: nitrate transport ATP-binding subunits C
and D
TIGR01186:proV: glycine betaine/L-proline transport ATP
binding subunit
TIGR01187:pocA: spermidine/putrescine ABC transporter
ATP-binding subunit
TIGR01188:drdA: daunorubicin resistance ABC transporter
ATP-binding subunit
TIGR01189:ccmA: heme exporter protein CcmA"
/codon_start=1
/transl_table=11
/product="ABC-type transporter, ATPase component"
/protein_id="BAB98339.1"
/db_xref="GI:21323713"
/translation="MTTTPAIDVTLVRYDYTAVKGLNFHVRQVGFVGLLTNGAG
KSTLEVIKGLSAPSSGTVIRISGLDPVADRAILRPELQMLQSGELSQLTVAETMDM
WNGTTPRATKDVADVLLHRENVKVGALSGQERLDLACALLGDSILFLDEPT
GLDPSRRHWQLLLDLKQGVTHMLTTHYLEAEFLCDRIAMNAGEIAVEGTLDE
LVAREKSIISFVLRGGQVELPVLGABIIIRDNNHVRITATTLQOHTLEILTWAAETGI
ALEGFAKPATLESVFMDSLENTSLQTA"
4003..4836
/gene="Cg10947"
4003..4836
/gene="Cg10947"
/codon_start=1
/transl_table=11
/product="ABC-type transporter, permease components"
/protein_id="BAB98340.1"
/db_xref="GI:21323714"
/translation="MTTSTARGLQHAHPERTKTSLEKAEWLQPRNKLTLLFM
ATVPVGPILLFLIGNGAASNSDFVMTLLVQVITVLSMATTRRDEVLKR
LRTGEARDIDIGALIFPGALITLFTVVIPLMLVGLGAPAPINLVPALVILGLLL
CSALAMTSGTFRNAEAQMTSPVFMALNGILSGISIRFVFGSDIADILATPFAAIS
DLVQIGWAGATFADSVGVGEANFAGIFQDMLPLIGLILAAATAAANAANRYRWDSY
R"
4879..6138
/gene="Cg10948"
4879..6138
/gene="Cg10948"
/EC_number="2.7.3.-"
/codon_start=1
/transl_table=11
/product="Two-component system, sensory transduction
```

histidine kinases"
/protein_id="BAB98341.1"
/db_xref="GI:21323715"
/translation="MNKQFTAGTARFSGVSLASPVTPAPNSWRRLNPKAYTL
VTRLSQAI PVLLSLAYFLIGVANAGTLPSPFVLLGFSVILLITVLVYEPQSLNS
HRRSVQPFPTGLVNLGVVSVLQIPGLNMSDNTRATLITLTCVFLSLAYI
PMNTRVWLMASVAVMTSTTIDYLSALWVPIPLMAGTVRLSVMTVDVWKEVER
RELEASVTEERLPAQELHDTLQGLHAMSVMKSELALAKRGDORLENELEOK
LRTSMSEMDVSVGYRTVNLATIEGASLADAHILHSVIGTTSQVSPARELCAM
LVREATTNILRHSDATDPTLSTSEVRMDNNGVNDIGRLSGLSALRSRBSAGMTL
IVSREDDQFVRMLINAPANTPAEKA"

6139..6750
/gene="Cg10949"
6139..6750
/gene="Cg10949"
/note="PF00072:Response regulator receiver domain
PF00196:Bacterial regulatory proteins, luxR family"
/codon_start=1
/transl_table=11
/product="two-component system, response regulators
consisting of a CheF-like receiver domain and a HTH
DNA-binding domain"
/protein_id="BAB98342.1"
/db_xref="GI:21323716"
/translation="MISISIADEALIASLALLSLPELDLVRPTAGSGELIETWA
DPSNTDVCVLDLQGLGIDITATRLNETTPDLAVLIVTSHAPROLKALAAVGLG
FLPKSTADEFATIRTVHAGRRYIDPELAAMTISAGESPLNREERVELAGQSLA
EELAVAAHAPGTITRNLYSQAMTKVGNRFEAPTRARELGM"
complement(6781..6987)
/gene="Cg10950"
complement(6781..6987)
/gene="Cg10950"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB98343.1"
/db_xref="GI:21323717"
/translation="MKHYSMYFAGAPLTPHTKEVWFMDIVSIWNHFWBSWNNVN
MTGLPRETFVPFVGVSTVLAVLSF"

7048..8205
/gene="Cg10951"
7048..8205
/gene="Cg10951"
/codon_start=1
/transl_table=11
/product="tRNA-processing ribonuclease BN"
/protein_id="BAB98344.1"
/db_xref="GI:21323718"
/translation="MTAWLCSFSGVDYDMNGVQPOEHLDTLIAADPHNGPEN
SGDKERLNPQMKYALNRTVRDVPDGLDLAALLTFSSILSIAPAVLLGYSVITIF
LASDTEILNVRDVNGYVPEQSHVNGVIDSIAGSAAAGVGVAVGVTALTWTSF
AYVAFSRCANAVYGRSGRTLIKRWMLLEFLNALLGIIILVSWLNELTMGIF
APIAEPHLNLTFLDREMPIIWIWVRFPIVGVLLMFVAILIYWNAPRKFWRL
SLGSFLAIVGILLAGVGLNFTFLFAAFSSVAGVSGVLLAVFTALVFNICLIIGLKID
VEISRAKQLQAGPAEDYSLVPPRSIEKVKMKQKQORLMDQAAAREEN"

8672..10114
/gene="Cg10952"
8672..10114
/gene="Cg10952"
/note="PF00005:ABC transporter
PF00664:ABC transporter transmembrane region."
/codon_start=1
/transl_table=11
/product="ABC-type multidrug/protein/lipid transport
system, ATPase component"
/protein_id="BAB98345.1"
/db_xref="GI:21323719"
/translation="MNRTHASSILRMIRRGKAVGAFFLGMWOLSEALVPIAI
GLVHVALVTDRLRLVGLVAFVVLVFLVLSYRFGSRALNRAVNFESHALRVEVAD
HALKNDPRMLVPEVMSKSTADSSSTRIFGIGTGSATGELGAATYLLISDWLV
GLVLIIVPIISGVWASAKGISKRSVTOEKLEASQASDMMGLIRVKAIGGERH
AVKTEKASQASAAVDTAVASGVKVIAGELSIATVNLAAVLLAGRVVTGELGPQ
LTAIVGVAVVLSRPILLNSINASATAGMAERVAENFLNDESQAQYESSTINDGE

FLVIVPASTPLPHGNILATPHAADIEFTLRNISMHEHNVVIDPQVIRASGLTDI
IEVDGLDAPVDTGSSNLGGQRVALRALHADAEVLVIMDPTSAVDSVTEVSTAQ
IKQLRAGKATVIVVSSPAFYNLADRVISHV"

10107..11843
/gene="Cg10953"
10107..11843
/gene="Cg10953"

Query Match 100.0%; Score 711; DB 1; Length 333150;
Best Local Similarity 100.0%; Pred. No. 5.9e-169;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGATCATGAAATCTTCATACAGTCTGCTTTGGGGCCAGTCTTTTACTCTCC 60
Db 329003 ATGTGATCATGAAATCTTCATACAGTCTGCTTTGGGGCCAGTCTTTTACTCTCC 328944

Qy 61 ATCCGACCGCAGAACTGCTGATTAACAAGGAATTAAGCGCGAAGGACTCATTTGG 120
Db 328943 ATCCGACCGCAGAACTGCTGATTAACAAGGAATTAAGCGCGAAGGACTCATTTGG 328884

Qy 121 GTTCTTCTGCTGTGTTTAAATTTCTGACGCTTTTGTTCATCGCCGCGACCTTGGGGGTT 180
Db 328883 GTTCTTCTGCTGTGTTTAAATTTCTGACGCTTTTGTTCATCGCCGCGACCTTGGGGGTT 328824

Qy 181 GATCTTTTGTCCAAATGCGCCGCGATGCTCGATATTAAGCGCTGGGCTGGCATCGCT 240
Db 328823 GATCTTTTGTCCAAATGCGCCGCGATGCTCGATATTAAGCGCTGGGCTGGCATCGCT 328764

Qy 241 TACCTGTTATGTTTTCGCTCATCGCAGCAAGAGCGCATGACAAACAGAGTGGAGCG 300
Db 328763 TACCTGTTATGTTTTCGCTCATCGCAGCAAGAGCGCATGACAAACAGAGTGGAGCG 328704

Qy 301 CCACAGATCATTTGAGAAACAGAACCAACCGTCCCGATGACAGCCCTTTGGGGCGTTTCG 360
Db 328703 CCACAGATCATTTGAGAAACAGAACCAACCGTCCCGATGACAGCCCTTTGGGGCGTTTCG 328644

Qy 361 GCGGTGGCCACTGACAGCGGCAACCGGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 420
Db 328643 GCGGTGGCCACTGACAGCGGCAACCGGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 328584

Qy 421 GTTTCGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGTTGAAACCGCAATCGGTAT 480
Db 328583 GTTTCGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGTTGAAACCGCAATCGGTAT 328524

Qy 481 TTGACAGCGCTTTGTGTTATTCGCGCGGCTCGCGCGCAATACGCGACCGAGCGGTGG 540
Db 328523 TTGACAGCGCTTTGTGTTATTCGCGCGGCTCGCGCGCAATACGCGACCGAGCGGTGG 328464

Qy 541 ATTTTCGCGCTGGCGGCTTCGCGCAAGCTGATCTGCTCCCGCTGGTGGTTCGCGC 600
Db 328463 ATTTTCGCGCTGGCGGCTTCGCGCAAGCTGATCTGCTCCCGCTGGTGGTTCGCGC 328404

Qy 601 GCACGACGATGTCACGCGCGCTGTCACGCCCAAGCTGCGCGCTGGATCAACGCTGTC 660
Db 328403 GCACGACGATGTCACGCGCGCTGTCACGCCCAAGCTGCGCGCTGGATCAACGCTGTC 328344

Qy 661 GTGCGAGTTGTGATGACCGCATTCGCGCATCAACTGATCTGATGGGTTAG 711
Db 328343 GTGCGAGTTGTGATGACCGCATTCGCGCATCAACTGATCTGATGGGTTAG 328293

RESULT 7
AX127147/c
LOCUS AX127147 349980 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 7063 from Patent EP1108790.
ACCESSION AX127147 AX114121
VERSION AX127147.1 GI:14041135
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.
REFERENCE 1

AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7063 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES Location/Qualifiers
source 1..349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3.309.400) split in 11, seq 7063
1.200.001 1.549.980"
ORIGIN
Query Match 100.0%; Score 711; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 5.9e-169;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGTGATCATGGAATCTTCATTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTGCC 60
Db 128953 ATGGTGATCATGGAATCTTCATTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTGCC 128894
QY 61 ATCGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGGAAGACTCATTTGCG 120
Db 128893 ATCGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGGAAGACTCATTTGCG 128834
QY 121 GTTCTTCGTGCTGTTAAATTTCTGACCTCTTTTGTTCATCGCGGACCTTTGGCGTT 180
Db 128833 GTTCTTCGTGCTGTTAAATTTCTGACCTCTTTTGTTCATCGCGGACCTTTGGCGTT 128774
QY 181 GATCTTTTGTCAATGCGCGCGGATCGTCTGATATTATGCGTGGGCTGATCGCT 240
Db 128773 GATCTTTTGTCAATGCGCGCGGATCGTCTGATATTATGCGTGGGCTGATCGCT 128714
QY 241 TACCTGTTATGTTTTCGCTCATGCGCAGAAAGACCCATGACAAACAAGGTGGAAGCG 300
Db 128713 TACCTGTTATGTTTTCGCTCATGCGCAGAAAGACCCATGACAAACAAGGTGGAAGCG 128654
QY 301 CCACAGATCATTTGAAGAAACAGAACCCAGTCCGATGACAGCGG 420
Db 128653 CCACAGATCATTTGAAGAAACAGAACCCAGTCCGATGACAGCGG 128594
QY 361 GCGGTGGCCACTGACAGCGCAACCGGCTGCGGTGAGGTGAGTGCATTAAGCAGCGG 420
Db 128593 GCGGTGGCCACTGACAGCGCAACCGGCTGCGGTGAGGTGAGTGCATTAAGCAGCGG 128534
QY 421 GTTTGGGTAAGCCCATGTTGATGGAATCGTCTGATCTGCTGATGATGATGATGATGAT 480
Db 128533 GTTTGGGTAAGCCCATGTTGATGGAATCGTCTGATCTGCTGATGATGATGATGATGAT 128474
QY 481 TTGGACGCGTTTGTGTTTATCGGCGCGTCCGCGCAATACGCGACACCGGAGCGTGG 540
Db 128473 TTGGACGCGTTTGTGTTTATCGGCGCGTCCGCGCAATACGCGACACCGGAGCGTGG 128414
QY 541 ATTTTCGCGCTGGCGGTTTCGCGCAAGCCCTGATCTGGTTCCCGCTGCTGGGTTTCGCG 600
Db 128413 ATTTTCGCGCTGGCGGTTTCGCGCAAGCCCTGATCTGGTTCCCGCTGCTGGGTTTCGCG 128354
QY 601 GCGACGANTGTACGCGCGCTGTCCAGCCCGCAAGGTGTGGCGTGTGATCAACGCTGTC 660
Db 128353 GCGACGANTGTACGCGCGCTGTCCAGCCCGCAAGGTGTGGCGTGTGATCAACGCTGTC 128294
QY 661 GTGGCAGTTGTGATGACCGGATTTGGCCATCAAACTGATGTTGATGGTTAG 711
Db 128293 GTGGCAGTTGTGATGACCGGATTTGGCCATCAAACTGATGTTGATGGTTAG 128243
RESULT 8
AX123539
LOCUS AX123539 708 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 3455 from Patent EP1108790.
ACCESSION AX123539
VERSION AX123539.1 GI:14041027

KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium
REFERENCE 1
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 3455 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES Location/Qualifiers
source 1..708
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
ORIGIN
Query Match 99.6%; Score 708; DB 6; Length 708;
Best Local Similarity 100.0%; Pred. No. 1.5e-168;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTGATCATGGAATCTTCATTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTGCC 60
Db 1 ATGCTGATCATGGAATCTTCATTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTGCC 60
QY 61 ATCGGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGGAAGACTCATTTGCG 120
Db 61 ATCGGACCGCAGATGTACTGTGATTAAACAAGGAATTAAGCGGAAGACTCATTTGCG 120
QY 121 GTTCTTCGTGCTGTTAAATTTCTGACCTCTTTTGTTCATCGCGGACCTTTGGCGTT 180
Db 121 GTTCTTCGTGCTGTTAAATTTCTGACCTCTTTTGTTCATCGCGGACCTTTGGCGTT 180
QY 181 GATCTTTTGTCAATGCGCGCGGATCGTCTGATATTATGCGTGGGCTGATCGCT 240
Db 181 GATCTTTTGTCAATGCGCGCGGATCGTCTGATATTATGCGTGGGCTGATCGCT 240
QY 241 TACCTGTTATGTTTTCGCTCATGCGCAGAAAGACCCATGACAAACAAGGTGGAAGCG 300
Db 241 TACCTGTTATGTTTTCGCTCATGCGCAGAAAGACCCATGACAAACAAGGTGGAAGCG 300
QY 301 CCACAGATCATTTGAAGAAACAGAACCCAGTCCGATGACAGCGG 360
Db 301 CCACAGATCATTTGAAGAAACAGAACCCAGTCCGATGACAGCGG 360
QY 361 GCGGTGGCCACTGACAGCGCAACCGGCTGCGGTGAGGTGAGTGCATTAAGCAGCGG 420
Db 361 GCGGTGGCCACTGACAGCGCAACCGGCTGCGGTGAGGTGAGTGCATTAAGCAGCGG 420
QY 421 GTTTGGGTAAGCCCATGTTGATGGAATCGTCTGATCTGGTTCCCGCTGCTGGGTTTCGCG 480
Db 421 GTTTGGGTAAGCCCATGTTGATGGAATCGTCTGATCTGGTTCCCGCTGCTGGGTTTCGCG 480
QY 481 TTGGACGCGTTTGTGTTTATCGGCGCGTCCGCGCAATACGCGACACCGGAGCGTGG 540
Db 481 TTGGACGCGTTTGTGTTTATCGGCGCGTCCGCGCAATACGCGACACCGGAGCGTGG 540
QY 541 ATTTTCGCGCTGGCGGTTTCGCGCAAGCCCTGATCTGGTTCCCGCTGCTGGGTTTCGCG 600
Db 541 ATTTTCGCGCTGGCGGTTTCGCGCAAGCCCTGATCTGGTTCCCGCTGCTGGGTTTCGCG 600
QY 601 GCACGAGCATTGTACGCGCGCTGTCCAGCCCGCAAGGTGTGGCGTGTGATCAACGCTGTC 660
Db 601 GCACGAGCATTGTACGCGCGCTGTCCAGCCCGCAAGGTGTGGCGTGTGATCAACGCTGTC 660
QY 661 GTGGCAGTTGTGATGACCGGATTTGGCCATCAAACTGATGTTGATGGTTAG 708
Db 661 GTGGCAGTTGTGATGACCGGATTTGGCCATCAAACTGATGTTGATGGTTAG 708
RESULT 9
BD165656

Db	361	GC	GTGCGCACTGACACCGGCAACCGCGGTGCGGTGAGCGTGCATTAAGCAGCGG	420			
Qy	421	GT	TTGGGTAAGCCCATGTTGATGGCAATCGTGTCTGACTGTGGTGAACCCGGAATGCGTAT	480			
Db	421	GT	TTGGGTAAGCCCATGTTGATGGCAATCGTGTCTGACTGTGGTGAACCCGGAATGCGTAT	480			
Qy	481	TT	TGACAGCGTTTGTGTTTATCGGCGCGTTCGGCGCGCAATACGGCGACACCGAGCGTGG	540			
Db	481	TT	TGACAGCGTTTGTGTTTATCGGCGCGTTCGGCGCGCAATACGGCGACACCGAGCGTGG	540			
Qy	541	ATT	TTTTCGCGCGTTCGGCGCGTTCGGCGCAACCTCGATCTGTGTTCCCGTGGTGGGTTTCGCGC	600			
Db	541	ATT	TTTTCGCGCGTTCGGCGCGTTCGGCGCAACCTCGATCTGTGTTCCCGTGGTGGGTTTCGCGC	600			
Qy	601	GC	AGCAGCATGTCACGCGCGCGTTCGACGCCCAAGGTGTGGCGCTGTGATCAACGTCGTC	660			
Db	601	GC	AGCAGCATGTCACGCGCGCGTTCGACGCCCAAGGTGTGGCGCTGTGATCAACGTCGTC	660			
Qy	661	GT	GCGCAGTTGTGATGACCGCATTCGGCCATCAAACTGATGTTGATCGGT	708			
Db	661	GT	GCGCAGTTGTGATGACCGCATTCGGCCATCAAACTGATGTTGATCGGT	708			
RESULT 10							
AX643030							
LOCUS	AX643030		712 bp	DNA linear PAT 24-FEB-2003			
DEFINITION	Sequence 9 from Patent EP1266966.						
ACCESSION	AX643030						
VERSION	AX643030.1	GI:28550160					
KEYWORDS	Corynebacterium glutamicum						
SOURCE	Corynebacterium glutamicum						
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.						
REFERENCE	1. Gunji, Y. and Yasueda, H.						
AUTHORS	Method for producing l-lysine or l-arginine by using methanol						
TITLE	assimilating bacterium						
JOURNAL	Patent: EP 1266966-A 9 18-DEC-2002;						
FEATURES	Ajinomoto Co., Inc. (JP)						
source	Location/Qualifiers						
	1..712						
	/organism="Corynebacterium glutamicum"						
	/mol_type="unassigned DNA"						
	/db_xref="taxon:1718"						
CDS	1..375						
	/note="unnamed protein product"						
	/codon_start=1						
	/transl_table=11						
	/protein_id="CAD67812.1"						
	/db_xref="GI:28550161"						
	/translation="WMVMEIFITGLLGLSLLSIGPQNVLVKIQIKRGLLIIVLVN						
	CLISDFVLEFAGTGLVLDENNAFVILDIIMRWGGIAYLLWFAVAAKADANTNKVAPD						
	IIEETEPTVPDDTPLGVFGGSH"						
ORIGIN							
Query Match	98.0%; Score 696.8; DB 6; Length 712;						
Best Local Similarity	99.6%; Pred. No. 1.1e-165;						
Matches	709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;						
Qy	1	ATG	TGATCATCGAAATCTTCATTACAGGTCTGCTTTTGGGGGCCAGTCTTTTACTGTCC	60			
Db	1	ATG	TGATCATCGAAATCTTCATTACAGGTCTGCTTTTGGGGGCCAGTCTTTTACTGTCC	60			
Qy	61	AT	CGGACCGCAGAAATGATCTGTGATTAACAAGGAATTAAGCGGAGGACTCATTTGGC	120			
Db	61	AT	CGGACCGCAGAAATGATCTGTGATTAACAAGGAATTAAGCGGAGGACTCATTTGGC	120			
Qy	121	GTT	CTCTCTGTGTTTAATTTCTTCAGTCTTTTGTTCATCGCGCGCACCTTGGCGGTT	180			
Db	121	GTT	CTCTCTGTGTTTAATTTCTTCAGTCTTTTGTTCATCGCGCGCACCTTGGCGGTT	180			

181 GATCTTTGTCCTAATCGCGCGCATGCTGCTGATATTATGCTGCTGGGTGGCATGCT 240
 181 GATCTTTGTCCTAATCGCGCGCATGCTGCTGATATTATGCTGCTGGGTGGCATGCT 240
 241 TACCTGTTATGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 241 TACCTGTTATGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 301 CCACAGATCATGAAAGACAGACACACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 359
 301 CCACAGATCATGAAAGACAGACACACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 360 GCGGTGCGCACTGACACACCGCAACCGCGTGGGTGGGTGGGTGGGTGGGTGGGT 419
 361 GCGGTGCGCACTGACACACCGCAACCGCGTGGGTGGGTGGGTGGGTGGGTGGGT 420
 420 GGTGTTGGTAAAGCCCATGTTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
 421 GGTGTTGGTAAAGCCCATGTTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 480 TTTGGACCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 481 TTTGGACCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 540 GATTTTCGCGCTGCGCGCTGCTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
 541 GATTTTCGCGCTGCGCGCTGCTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 600 CGACGACATGTCAGCGCGCTGCTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 601 CGACGACATGTCAGCGCGCTGCTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 660 CTGTCAGCTGCTGATGACCGCATGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCT 711
 661 CTGTCAGCTGCTGATGACCGCATGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCT 712

RESULT 11
 E54483
 LOCUS 1568 bp DNA linear PAT 27-AUG-2002
 DEFINITION Heat-resistant lysine biosynthesis enzyme gene of thermophilic
 coryneform bacterium.
 ACCESSION E54483
 VERSION 1
 KEYWORDS JP 2001120270-A/7.
 SOURCE Corynebacterium thermoaminogenes
 ORGANISM Corynebacterium thermoaminogenes
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriidae; Corynebacterium.
 REFERENCE 1 (bases 1 to 1568)
 AUTHORS Itaya, M., Kimura, E., Kawara, Y. and Sugimoto, S.
 TITLE Heat-resistant lysine biosynthesis enzyme gene of thermophilic
 coryneform bacterium
 JOURNAL Patent: JP 2001120270-A 7 08-MAY-2001;
 COMMENT AUINOMOTO CO INC
 OS Corynebacterium thermoaminogenes
 PN JP 2001120270-A/7
 PD 08-MAY-2001
 PF 01-NOV-1999 JP 1999311148
 PI MINOKU ITAYA, EICHIRO KIMURA, YOSHIO KAWARA, SHINTACHI SUGIMOTO PC
 C12N15/09/(C12N15/09, C12N15/00, C12N15/00, C12N15/00, C12N15/00) CC
 FH Key Location/Qualifiers
 FT CDS (778) ..(1470).
 FEATURES Location/Qualifiers
 source 1..1568
 /organism="Corynebacterium thermoaminogenes"
 /mol_type="genomic DNA"
 /db_xref="taxon:161484"
 ORIGIN
 Query Match 44.78; Score 317.6; DB 6; Length 1568;
 Best Local Similarity 67.74; Pred. No. 1.5e-69;
 Matches 481; Conservative 0; Mismatches 214; Indels 15; Gaps 2;

1 ATGCTGATCATGGAATCTTATACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60
 778 ATCGGGACATGGAATCTTGTGTACCGGTTTGTGTGGAGCCAGTCTGTGTGGCC 837
 61 ATCGGACCGCAGAAATGATCTGTGATTAACAAGGAATTAAGCGGAGGACTCATTTGG 120
 838 ATCGGCGCACAGAAATGCTGTGATCAACAGGGATCAAAACGGAGGGCATCAGGCC 897
 121 GTTCTTCTGCTGTTTAAATTTCTGAGCTCTTTTGTTCATCGCGGACCTTTGGGCGTT 180
 898 GTCATCATGCTCTGTCTGTGCTGCGAGCTGCTGTTCACCTTCGCGCACCTTCGGGGTC 957
 181 GATCTTTGCTCAATGCGCGCGCATGCTGCTGATATTATGCTGCGGTGGCATCGCT 240
 958 GGCCTGATCTCGACACCGCGCATCTTCTGACATCTGCTGCTGCTGCTGCTGCTGCT 1017
 241 TACCTGTTATGCTTTGCGCTGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 1018 TACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
 301 CCACAGATCATGGAAGAAACAGAACCAACCGTCCCGATGACACGCTTTGGGGGCTCG 360
 1078 ACCTTTGCT-----GAGCATTCGAAACCGGTTG---CGGACGCTCGGCTCC 1122
 361 GCGGTGGCCACTGACACGCGCAACCGGCTGCGGTGGAGTGGAGTGGAGTGGAGTGGAG 420
 1123 GCGGGGGGCTGACGAGAAACAGACCGCGCTCGCATCACATCAGGACCCCGGAG 1182
 421 GTTGGGTAAGCCCATGTTGATGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 1183 GTCTGGGTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1242
 481 TTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 1243 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1302
 541 ATTTTCGCGCTGCGGCTTTCGCGCAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 1303 ATCTTCGCTGCGGCTGCTTTCGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362
 601 GCGAGCAGATGTCACGCGCGCTGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
 1363 GCGGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422
 661 GTGCGAGTGTGATGACCGCATGCGCATCAACTGATGCTGCTGCTGCTGCTGCTGCT 710
 1423 GTGGCGGTGGTGGTCCCGGATGGCGTGAAGCTGATCTGCTGCTGCTGCTGCTGCTGCT 1472

AB083133 1771 bp DNA linear BCT 06-APR-2002
 DEFINITION Corynebacterium efficiens lysG, lysE genes for Lysine export
 transcriptional regulatory protein, Lysine exporter protein,
 complete cds.
 ACCESSION AB083133
 VERSION 1
 KEYWORDS Corynebacterium efficiens
 SOURCE Corynebacterium efficiens
 ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriidae; Corynebacterium.
 REFERENCE 1
 AUTHORS Itaya, H., Kimura, E., Kawahara, Y. and Sugimoto, S.
 TITLE lysG, lysE of Corynebacterium efficiens
 JOURNAL Published Only in Database (2002)
 REFERENCE 2 (bases 1 to 1771)
 AUTHORS Itaya, H., Kimura, E., Kawahara, Y. and Sugimoto, S.
 TITLE Direct Submission
 JOURNAL Submitted (04-APR-2002) Hiroshi Itaya, AJINOMOTO CO., INC.
 Fermentation & Biotechnology Laboratories, 1-1, Suzuki-cho,
 Kawasaki-Ku, Kawasaki, Kanagawa 210-8681, Japan

(E-mail:hiroshi_itaya@ajinomoto.com, Tel:81-44-244-7123 (ex.4146),
Fax:81-44-222-0129)

```

FEATURES
    Location/Qualifiers
        1..1771
            /organism="Corynebacterium efficiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:152794"
            complement(47..919)
            /gene="lysG"
            complement(47..919)
            /gene="lysG"
            /codon_start=1
            /transl_table=11
            /product="lysine export transcriptional regulatory protein"
            protein_id="BAB88826.1"
            /db_xref="GI:20065739"
            translation="MNPILHLLTIIIDEGSPENASIALSISPSAVSORIKALEKSVG
            RVLSRTOPAVATEAGEVLVQARFMALLOAETREGLAERLDEIPLTVAINADSLSTM
            FPPVFAVHGAVATLRLVEDEATLSLRGSLVLCATVREADPVGCEVLRIGVMR
            HLPVATPELRARYTVDDQDWMFVLRFGNDVLDGRVDFDGVAVARRVSVVPS
            AFGCEAVRLGLGLPEAQAAPLAAGDVQDLEKVVDTPLVQWRLESLRLARL
            TDAAVDAARAGLRT"
            981..1676
            /gene="lysE"
            981..1676
            /gene="lysE"
            /codon_start=1
            /transl_table=11
            /product="lysine exporter protein"
            protein_id="BAB88827.1"
            /db_xref="GI:20065740"
            translation="WRDMEIFVTGLLGASLLAIGPQNVLVKQIKRGITAVLIIV
            CLISDVLTGTLGVLGSDYTDPIILDLRWGCIAYLWFAVMAARDALRAKTEVIF
            VESPEVAASAGSGVYTKQRPRIITSTGRQVYRPMALVLTWLNPNYALDAFV
            FIGGVAOVGTGTRWIPAGATAAASLVWFLVGVGAALSRPLSPRVWRWINGVAV
            VLTGLAVKLLWG"

ORIGIN
    Query Match 44.7%; Score 317.6; DB 1; Length 1771;
    Best Local Similarity 67.7%; Pred. No. 1.5e-69;
    Matches 481; Conservative 0; Mismatches 214; Indels 15; Gaps 2;

QY 1 ATGGTGATCATGGAATCTTCAATACAGGTCTGCTTTGGGGCCAGTCTTTTACTGTC 60
DB 981 ATGGGGACATGGAATCTTGTACCGGTTGTTGTTGGAGCCAGTCTGCTGTGGCC 1040
QY 61 ATCGGACCGCAGAAATGACTGTGATTAACAAGGAATTAAGCGCGAGGACTCATTCGG 120
DB 1041 ATCGGCCCAAGAAATCTCTGTGATCAACAGGGCATCAACCGCGAGGCATCAGGCC 1100
QY 121 GTTCTTCGTGTGTTAAATTTCTGACGTCTTTTGTTCATCGCCGGCACCTTGGGGTT 180
DB 1101 GTCATCATCTGTCTGTCTGTCGAGCTGGTCTGTTTCACTCCGCGCACCTCGGGGTC 1160
QY 181 GATCTTTTGTCCATCGCGCGCATCGTCTCGATATTATGCGTGGGGTGGCATCGCT 240
DB 1161 GGCCTGATCTCCGACACCCCGGATCATTTCTCGACATCTGCGTGGTGGCGCATCGCC 1220
QY 241 TACCTGTTATGTTTCGCTCATGCGCAGCGAAGACGCATGACAAACAAGGTGGAAGCG 300
DB 1221 TACCTGCTGTGGTTGCGGTGATGCGCGCGCGCGCCCTCGCGCCCGCACCGAGTA 1280
QY 301 CCACAGATCATTTAGAGAAACAGAACACACACGTCGCCGATGACAGCCCTTGGCGGTTCG 360
DB 1281 ACCTTTGTC-----GAGCATTCGGAACCCGTTG---CGGACAGCTCCGCCCTCC 1325
QY 361 GCGGTGGCCACTGACACGCGCAACCGGCTGCGGTGGAGGTGAGCTCGATAGCAGCGG 420
DB 1326 GCGGGGGCGGTGACGACGAAACACGACCGCGCTCCGATCATCATGCGCACCGCGGAG 1385
QY 421 GTTTGGGTAAAGCCCATGTTGATGGCAATCGTCTGCACTGCTGTTGAACCCGCAATCGPAT 480

```

```

Db 1386 GTCTGGTCCAGGCCCATCTCATGGCCATTGTGCTGACCTGGCTCAATCCCAATGCCCTAC 1445
QY 481 TTGAGCCGCTTTGTTATTCGGCGGCGTTCGGCGGCACATACCGCGACACCGGCGGTGG 540
Db 1446 CTGGATCCCTTCTCTTCATCGGTGGTGTCCGAGCCAGTACGGGGAGACCGGTCGGTGG 1505
QY 541 ATTTTTCGCGCTGCGCGCTTCGGCGCAACCTGATCTGCTTCCCGCTGGTGGGTTCGGCG 600
Db 1506 ATCTTCGCTGCGGTGCTTCCGCGCAGCTGTGCTTCCCACTGGTGGTTACGGC 1565
QY 601 GCACGACGATTTGTCACGCCCGCTGTCCAGCCCAAGAGTGTGGCGCTGGATCAAGTGTGC 660
Db 1566 GGGGCGGCACTGTGCGCTCCCTGTCCAGCGCGCGGTCTGGGCTGGATCAACATAGGT 1625
QY 661 GTGGCAGTTGTGATGACCGCATTTGGCNCATCAACTGATGTTGATGGGTTA 710
Db 1626 GTGGCCGCTGTGCTCACCGGATTTGGCGTGAAGCTGATCTGTGATGGGTTA 1675

RESULT 13
AP005218/c 308650 bp DNA linear BCT 10-JUL-2003
LOCUS Corynebacterium efficiens YS-314 DNA, complete genome, section 5/11.
ACCESSION AP005218 BA000035
VERSION AP005218.1 GI:23493016
KEYWORDS Corynebacterium efficiens YS-314
SOURCE Corynebacterium efficiens YS-314
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.
TITLE Comparative Complete Genome Sequence Analysis of the Amino Acid Replacements Responsible for the Thermotability of Corynebacterium efficiens
JOURNAL Genome Res. 13 (7), 1572-1579 (2003)
MEDLINE 22723752
PUBMED 12840036
REFERENCE 2 (bases 1 to 308650)
AUTHORS Kawarabayasi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066 Japan
COMMENT [E-mail: bio@nitech.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424] Kawarabayasi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan
Tharaki, 305-8566 Japan
Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
Itoh, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan
Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan
Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan
The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.
Location/Qualifiers
    1..308650
        /organism="Corynebacterium efficiens YS-314"
        /mol_type="genomic DNA"
        /strain="YS-314"
        /db_xref="taxon:196164"
        complement(83..241)
        /note="CE1178"
        /codon_start=1
        /transl_table=11
        /product="hypothetical protein"
    CDS

```

/protein_id="BAC17988.1"
/db_xref="GI:23493017"
/translation="MTTGLILGTGPDAGLAKLGSIDIPRGVGIHRFRKPPVALQVR
TGAGARD"
365..742
/note="CE1179"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17989.1"
/db_xref="GI:23493018"
/translation="WEKQQQSQKTOFASVDHLSDDAAAMFIDNELSRGAMHRAIRLH
LVHCAECRQIARQDITVDVFTCRNDNVTPADSLAKMLSLANECAPGCAADAAL
QRPESFVAKVESVVRVRTQGR"
774..1271
/note="CE1180, similar to AJ292256--1|CAC21724.1| percent
identity: 26 in 155 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17990.1"
/db_xref="GI:23493019"
/translation="MFSNVNGEILLFIVGLIVIGPERILPMLDKAAIIVAARTAI
NNAQNLDITGEEDEIRKPMQIAQIRMSGPKTAITRALFDGDDSFIDSPFKKIM
ADDTAGQARHQVAKTGTTPHEAEVVERPADRIDTQSPNTPKGNNDGKPPAGGGYT
WTDLI"
complement (1353..2480)
/note="CE1181, similar to AL390975--15|CAC01353.1| percent
identity: 63 in 370 aa"
/codon_start=1
/transl_table=11
/product="putative ATP-binding protein Mrp"
/protein_id="BAC17991.1"
/db_xref="GI:23493020"
/translation="MMSQVTSASVSALSRVEDPEIGKPTTELGMKVSVAIDGSDVQV
EYITIAACPMKTIIVSNEAAIKIEGVGDVQVTTDVMGDEQRRLVSLRGNTADP
VIFPAQSGTTRYVYVASKGVGKVSMTYNLATAAKGLSVGLDADITGHVSVPGM
LGSARPHQVDMIMPQAHGVMKISIAHFVEGNAPVVRGPMHLRAIQOFSVDFWG
LDLIDLLDPGTDIAITIAQLPNAELLIVTPQAAADVAERAGSISLQTNOKVA
GVENMSAMVPGSTMEVFGSGGKIAQRLSMLTGEDVPVLSGVPIDPSLRVGGDV
GNPTAISEPSPPTTAINAIDATLAKKKSIVGKTLGLGVR"
complement (2513..3271)
/note="CE1182"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17992.1"
/db_xref="GI:23493021"
/translation="MTTATPWADAFDHVGWDLFOILEVDHTERAVAHLOEYFATDPRT
GPMVTGSLFNRGLGGDAQPVRENIITRDLLSLLDVPVPGVQVQLLEPPEDAQSI
DLEWVTRDTAVELPYDLEQIRGLIASIPTDVLGASKAAEHLDAENLLMTALRRNG
RGPTRVSKLIARPKLPFIQDSALKQNTNSGVFENFHTVLRAEDKALSRHLKGI
RTEAADSGNSALADLSIRVFVDVVMADHKKI"
complement (3282..3836)
/note="CE1183, similar to AE007003-3|AAK45527.1| percent
identity: 57 in 159 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17993.1"
/db_xref="GI:23493022"
/translation="MADYSRSDLPVLGTRKRRFKFDDTVGAYAEKVARFPGTGOY
LFWQITFIWIMLNGGFWAMPDYPFILLNIAFSTQAAVAAPLILAAQNRQEDRK
STILADRAEETKADTELTETELAGVRLAVGHWTRYDLRHELDLRLGFLERLEAKL
DAVARNDPDLINAPTOCKVD"
complement (3829..5124)
/note="CE1184, similar to AE007003-4|AAK45528.1| percent
identity: 52 in 417 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17994.1"
/db_xref="GI:23493023"

/translation="MRGITRLVAGRLAGVMVRGPDITVIGRVDVVVTVIRSGQDSRV
LGLVVLNNKRIFLPLMLTAAIEPEKNLISGVSLSRSPKARTGELOVIGDIIGTK
VHTDDPELQHLGRVETIADVELEQTTRDWIISVALIGSRPKRPTLYIAPWSH
VHGITTAGQETITQAEIAGFEDMRADVAKELYGLPKFORDVAVAGELDDRADIL
QELSDRQAEIETIDIERAADVLEEMDDAADLLGSELPEDKAGVLELMDPESSAT
VRLMGFNPTVGMALMTEPLTMDPSTTVAERALARNPDLPTSLASLVFVVRPTAT
PTGKYLGCVLKQLKLEPPSLGVGILDPLPPLVANDSQETAARFATFATVNLVCGPVL
DENGHLIGAVAVDDLLDMLPEDWRDAGIRPGVKEETLG"
5264..5782
/note="CE1185, similar to AE007003-6|AAK45530.1| percent
identity: 60 in 155 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17995.1"
/db_xref="GI:23493024"
/translation="MSMQATSGGNNRAELGLKKRPEGVPVGSQTYTEERQAVDMLS
DNEFFVAELTIVGDLMEVETVGRLTWGRVLAGGAAGMLGFLGVLVGLVSLWF
SSLLTGLVMGVVFGVLAAPVYASRGKRDFTSATQIVAGRYDILCRDPRAREARDMI
ARNMGLGHPRQ"
5805..6500
/note="CE1186"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17996.1"
/db_xref="GI:23493025"
/translation="MTTGVKALIGLVAGSLIVAGSGVDNREFFEPHQVAEQK
ENLYTAVDDPYDLGHNVISLEYKVLNGRFPAVIRHENLADFPDLNMGSDSLVIG
CTGALLEVNPVLAGELEEEYLAAVEAGETDINSGEWRDFTDALIGSLPSRLDADP
SNAQCPGANGLPQNTVPIYRNPTFSRADREILNWSAGAITTSELQELBEAKELGRS
RELVKFVASKGG"
6660..7373
/note="CE1187, similar to AE007262-3|AAK65408.1| percent
identity: 32 in 235 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17997.1"
/db_xref="GI:23493026"
/translation="METILIGVILFVACVQGAIGFGLGMLAAPIALMRPOLLGLLI
LLAFGSLATWARDGEVMEVGVNSLLGRIPGSLVIGTWAVMPVAGLKIVLATAV
ILGTLSILVNSPGHGRNSFLAGSGGFLGTTAIGGPPALIMRSMSRERVGTLS
VCFVIGSAISALTGALGMHMLQALVYAPAVIAGVYFLSGVGVNKLNRRLIFGS
VTISLIGSMVITQAGIF"
7592..8149
/note="CE1188"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17998.1"
/db_xref="GI:23493027"
/translation="MSRGDGFPAEPFPAADPRCGRTIKMQMTKGAIASAAALLLV
GGGTGFMWDEVTQSDQAIAGNLLASADAGKWTLNGQDVVDNNAVAVVGGDTLV
SGTQISYTSQDNLSATLITLTPGTSIGATADAALATELNDITITDILGELPSTIG
FNQTYTFRGVQSGVSALSIAKQA"
complement (8190..9443)
/note="CE1189, similar to AF024666-23|AAG03374.1| percent
identity: 84 in 408 aa"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC17999.1"
/db_xref="GI:23493028"
/translation="MAAGPYSIDPTTYLDELLAASPDLMREMQGFINOILSTQADQ
VGDYATVSDRVNTRNGRHRPDRVGTIDVAVPKLRTGSPFDDMLERTRAEER
AUTVTIATLKGVSTRMNDVATLGNLNSKQVSQVMDQDQMWEEFRPRLDTG
PLYVSCDALTMKVRGEVVKTSVLLATGVAAGYRELLGKQVATAESVETPDTG
DLKARGLENYLVSADHLGIQHAIGVFPNSWQRCTHFSKNLYGMWSKTQWPTLS
AMFHTIFQOPDAQSVMGARVAVFCQKFPDVADYLEALDELLAFATKSVMTKV
WSNPTELRNEIRRTDVVGLFNRDAVRLVAGVLAQHDHDDMIQOKRYSLSLEQ
TRAMTANIISDDITSEILIKDA"
complement (11439..15143)

gene


```

/misc_feature
/locus_tag="DIP1059"
/notes="HMPfam hit to PF00005, ABC transporter"
complement(579..797)
/locus_tag="DIP1059"
/notes="ProfileScan hit to P850100, 2nd half motif for
nucleotide binding, associated with p-loop."
complement(753..797)
/locus_tag="DIP1059"
/notes="ScanRegExp hit to PS00211, ABC transporters family
signature."
complement(1047..1124)
/misc_feature
/locus_tag="DIP1059"
/notes="ProfileScan hit to P850101, P-loop nucleotide
binding motif (does not find all)."
complement(1086..1109)
/misc_feature
/locus_tag="DIP1059"
/notes="ScanRegExp hit to PS00017, ATP/GTP-binding site
motif A (p-loop)."
complement(1223..2182)
/misc_feature
/locus_tag="DIP1060"
complement(1223..2182)
/locus_tag="DIP1060"
/notes="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system transmembrane component
sc151.26C TR:Q9S214 (EMBL:AL109848) (375 aa) fasta scores:
E(): 5e-40, 40.74% id in 324 aa, and to Escherichia coli
ferric enterobactin transport system permease protein fepg
fepg or B0589 SW:FEFG_ECOLI (P23877) (330 aa) fasta
scores: E(): 1.8e-45, 44.09% id in 322 aa"
/codon_start=1
/transl_table=11
/product="Putative iron-siderophore uptake system
transmembrane component"
/protein_id="CAE49582.1"
/db_xref="GI:38199915"
/translation="MRALVILSLILVTAAYSLVPIGAGLSTWQLTSETPLAHVVM
QWRLPVTVGLVGAALAIAGSLFQSLTRNPLGSPDIFGFGAYTVIAAFLGWSG
FGATVIGALIGSLVGVVVALSVTRIDGLRIILVGLIGSAMLNLRMLITRGELD
TALASAGAGSLGLRWTAIPACCVLAITTTIPRFDLVLSLGDILAVGLGLR
LQLTKLLVGGVFLVAVTAVAGPIAFVALASPHIARALTSSARTPLVEISVIGALL
IVADLIGRLRFPQLPVGVLVTVTGGMILNLIWLFARGSKENS"
/locus_tag="DIP1060"
/notes="Signal peptide predicted for DIP1060 by SignalP 2.0
HMM (Signal peptide probability 0.966) with cleavage site
probability 0.180 between residues 25 and 26;
signal-peptide site"
complement(1244..2116)
/misc_feature
/locus_tag="DIP1060"
/notes="HMPfam hit to PF01032, FecCD transport family"
order(complement(2105..2161), complement(1979..2044),
complement(1883..1939), complement(1805..1870),
complement(1718..1783), complement(1607..1657),
complement(1445..1510), complement(1319..1384),
complement(1247..1303))
/locus_tag="DIP1060"
/notes="9 probable transmembrane helices predicted for
DIP1060 by TMHMM2.0"
complement(1250..1564)
/misc_feature
/locus_tag="DIP1060"
/notes="BlastProDom hit to PD001557, PD001557"
complement(2184..3191)
/misc_feature
/locus_tag="DIP1061"
complement(2184..3191)
/locus_tag="DIP1061"
/notes="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system transmembrane component
sc151.27C TR:Q9S213 (EMBL:AL109848) (348 aa) fasta scores:
E(): 8e-36, 36.53% id in 323 aa, and to Escherichia coli
ferric enterobactin transport system permease protein fepd
or B0590 SW:FEFD_ECOLI (P23876) (334 aa) fasta scores:
E(): 5.2e-40, 40.37% id in 322 aa"
/codon_start=1

```

```

/transl_table=11
/product="Putative iron-siderophore uptake system
transmembrane component"
/protein_id="CAE49584.1"
/db_xref="GI:38199916"
/translation="MLRQPRVRLPLPVLITFLVILLCTIMSLGVGARNITSPDILA
ALRHYGDPNTSHIIGWRIIPRTLIATAAGASLSLAGVLIQALTRNPLADTVFGINAG
AAFTVIGALAGSLSHISIPALAILGAILAGACVVALSMNSCKGSDPLRLVLAGVAL
SAITGIGGLSLVNPQAPDELKSMWGNIDAGSYOPAAVAGVGLILGILVATCMRQ
LNALISDELATMGASIAKTRLTFTFIIVLAASATAAGVITFLGLMVPHIARNIV
GPNLLRLVATSLIGLFIIVLSADILGRILVPGEPFAGVAVAFGLFLAYATKRRK
I"
complement(3090..3191)
/misc_feature
/locus_tag="DIP1061"
/notes="Signal peptide predicted for DIP1061 by SignalP 2.0
HMM (Signal peptide probability 1.000) with cleavage site
probability 0.663 between residues 34 and 35;
signal-peptide site"
complement(2199..2525)
/misc_feature
/locus_tag="DIP1061"
/notes="BlastProDom hit to PD001557, PD001557"
complement(2205..3095)
/misc_feature
/locus_tag="DIP1061"
/notes="HMPfam hit to PF01032, FecCD transport family"
order(complement(3096..3161), complement(2834..2990),
complement(2838..2903), complement(2766..2831),
complement(2670..2735), complement(2544..2609),
complement(2394..2459), complement(2289..2354),
complement(2211..2276))
/locus_tag="DIP1061"
/notes="9 probable transmembrane helices predicted for
DIP1061 by TMHMM2.0"
3239..4282
/misc_feature
/locus_tag="DIP1062"
3239..4282
/locus_tag="DIP1062"
/notes="Similar to Escherichia coli iron(iii)
dicitrate-binding periplasmic protein precursor FecB or
B4290 SW:PECB_ECOLI (P15028) (300 aa) fasta scores: E():
Query Match 19.2%; Score 136.6; DB 1; Length 349535;
Best Local Similarity 52.4%; Pred. No. 2.2e-23;
Matches 354; Conservative 0; Mismatches 309; Indels 12; Gaps 2;
QY 2 TSGTGTATCATGAATCTTCATTACAGCTCTGCTTTTGGGGCCAGTCTTTTACTGTCCA 61
Db 37321 TGTGTATATAGTATGTCTATCTGCTGGAATTCCTCATGGGTCTGTCTTGTATCGTGGCA 37262
QY 62 TCGGACCGCAGAAATGTACTGGTGAATTAACAAGGAATTAAGCGCGAAGACTCATTTGCGG 121
Db 37261 TCGGGCCACAAAATGCCCTTGATTAATCGTCAAGGCATTAACGCGAAGGTCTAATCCCA 37202
QY 122 TTCTTCTGTGTGTTTAAATTCCTGAGCTTTTGTTCATCGCGGACCTTGGGGGTG 181
Db 37201 TTTGTGTATGATCACTCTCTGACGTGATTCCTCATTTTGGAGGCACTGCGGTGTGCG 37142
QY 182 ATCTTTTGTCCAAATCGCGCGCATGCTGCTCGCATATATATGCGCTGGGTGGCATCGCTT 241
Db 37141 GCGCATTAGTCGATCGAGCCCCATCGCACCTAGTAGTACTGAAATGGCTGGTGTAGCGT 37082
QY 242 ACTGTATATGTTTGGCTGTATGCGGCGAAGAACGCGATGACAAACAAGGTGGAGCGC 301
Db 37081 ACTTGTCTTAITTCGGATTCACTTGTTTTAAAGAAGCATTTAAACGTCACGGCAAGCGC 37022
QY 302 CACAGATCATTTGAAGAACACCAACCGTGGCCGATGACACGCTTTTGGCGGTTCGG 361
Db 37021 TCAGATC---GAGCAAGGACACCGGTGGCTTACGAAACAGTGGCTGACGATCATCCG 36965
QY 362 CGGTGGCCACTGACACCGCAACCGCGGTGCGGGTGGAGGTGAGCGTTCGATTAAGCAGCGG 421
Db 36964 GAGTGATCACTAAACACGCAAAAGCTCAGCCCAAGTCAGC-----GCAGCGCA 36914
QY 422 TTTGGGTAAAGCCCATGTTGATGGCANATCGTGTGACCTGTGTGACCCGANTGCGTATT 481

```

```

ncup://sequence.coucheux.mra.fr/K.Solanacearum.ntml.
Location/Qualifiers
1..190050
/organism="Ralstonia solanacearum"
/mol_type="genomic DNA"
/strain="GMI1000"
/db_xref="taxon:305"
complement(163..750)
/gene="RSC0382"
/note="synonym: RS03350"
complement(163..750)
/gene="RSC0382"
/function="miscellaneous; not classified regulator"
/note="Product confidence : probable
Gene name confidence : hypothetical
Predicted by Codon usage

```

predicted by Codon usage
predicted by Homology
predicted by Framed

/evidence-not experimental
/codon_start=1
/transl_table=11
/product="PROBABLE TRANSMEMBRANE PROTEIN"
/protein_id="CAD13913.1"
/db_xref="GI:17427395"
/db_xref="GOA:O8Y2F1"
/db_xref="SPTREMBL:O8Y2F1"

/translation="MTGPGESAPLIGALSGFGIGLASLIIVAGIAQNAVYLRQLRRE
YVGVLIICALCDMALIALGVAGMGTLSAHFAMLTAVRWAGAAFLIAYGARAPRAAW
RGAERLQANGKSHAOVLSALASALSLNPHVLDTVLLGALGGRVAMPANVAFAG
GAMCASTLAFSLILGFGARLLEPVFARPVAVRWLDALLIGAVWMAIALTLILMG"
complement(2924..3910)

gene

/gene="ftsY"
/note="synonyms: RS03354, RS0386"
complement(2924..3910)
/gene="ftsY"

CDS

/function="cell processes: cell division"
/note="Product confidence : putative"
Gene name confidence : putative

predicted by Codon usage
predicted by Homology
predicted by Framed

/codon_start=1
/evidence-not experimental

/transl_table=11
/product="PUTATIVE CELL DIVISION PROTEIN"
/protein_id="CAD13914.1"
/db_xref="GI:17427396"
/db_xref="GOA:O8Y2F0"
/db_xref="SPTREMBL:O8Y2F0"

/translation="NPDTTPAAPALDMQADDIEVTVPFVFEQARKGMNLSRLSGLSK
TSKSLTIFVGVDEALFBEELTALIMADAGVDATIELLEDRRVKAKRIETVAGGV
KTAURLLELLHLEKTLVLRDQPTVIMTAGVNGAGKTSIGLCKHFQYQGSVL
LAAGDTFPAAREQVLIWGNRVTVVAGSGDPAAVIFDANNAARARGIDIVNADTA
GRLPOLHMEELKKVRVIGKAMATAPHELTLLVIDANTGONALAQVKAQFDALGLTG
LIVTKLDGTAKGILAAIARQRPVVPVFIGVGEQVEDLPFSAREFADALLG"
4124..5629

gene

/gene="RSC0387"
/note="synonym: RS03355"
4124..5629
/gene="RSC0387"

CDS

/EC number="3.4.99.-"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : putative"

Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed

/codon_start=1
/evidence-not experimental

/transl_table=11
/product="PUTATIVE ZINC PROTEASE PROTEIN"
/protein_id="CAD13915.1"
/db_xref="GI:17427397"
/db_xref="GOA:O8Y2E9"
/db_xref="SPTREMBL:O8Y2E9"

/translation="MRSPQSVNSARTLRWAGITLAVALLAGEAQAQAPAPAAARQ
AAALTGAKVGTQADTHEYRLANGLSLIVKEDHRAPTVAHMVYHAGSIDENHGTGV
AHMLHMTKGFSTRVAGPGEFRRVAAALGRENAMTTRDFTMVFOQIEKSHLADVMALE
ADRMALNLTQDKFEPKMNVKKEERRMIDDSARATVYEQMLAVLNAAPVYRPTIGW
PSLDLMTVQDAQDMYHKVAPNNATVVTGVDNPDVEFQAQFTYKLOPHALPRY
AQDEPKQGVKRIWYKAPENPVVLYAYKAPPLRDVEKVDVPALEVLAVLDGYDNA
RLPNLLVTEKDEKGRADVNAGIDMNRGFSIFLLDGVDPADGHTTAEIEHQALRAQI
DRIARGDVTAEKRVAAQVAAQIYKEDSVFGQGMIEGMAETGLSWRDLDRILEKI
KSVTPAQVQVAKTYFTEDNLVATLLPOPIDPNKPAKVPQPMREGLR"
5629..6972

gene

/gene="RSC0388"
/note="synonym: RS03356"
5629..6972

CDS

/gene="RSC0388"
/EC number="3.4.99.-"
/function="miscellaneous; hypothetical/global homology"

Query Match 11.8%; Score 84.2; DB 1; Length 190050;

Best Local Similarity 47.9%; Pred. No. 3.9e-10;
Matches 242; Conservative 0; Mismatches 263; Indels 0; Gaps 0;

QY 207 CGTGTCTGATATATATGCGTGGGTGGATCGCTTACCTGTATGTGTTTGCCTGTCATGCG 266

Db 2423 CGTGTCTGCGCTGTGCTGATCTTGGCGCTGTGCGCATGCGGCTGATCGCGTGTGGCGCT 2482

QY 267 AGCGAAGACGCCATGATGAAACAAAGTGGAGGCCACAGATCATTTGAAGAACAGAAC 326

Db 2483 GCGGGCATGGGCACGCTGATCTCGGCCACCCCGGCATGCTGACCGCGGTGCGCTGGGC 2542

QY 327 AACGTTGCCCATGACACGCCCTTTTGGCGGTTTCGGCGGTGGCCACTGACACAGCGCAACCG 386

Db 2543 CGCGCGGCTTCTCTGCTGGCTTACGCGCACGCGGCTTCCGCGCGGCTGGCGCGCGCGC 2602

QY 387 GGTGCGGTGGAGGTGAGCGTGAATACAGCGGGTTTGGGTAAAGCCCATGTTGATGCG 446

Db 2603 CGAGCGGCTGACAGCGCGCAACGCGCATGCGAGGCTCGCATGCGCGGTGCTGGCGCTCCG 2662

QY 447 AATCGTGTGACCTGTGTAACCCGAATGCGTATTTGGACCGGTTTGTGTTTATCGCGCG 506

Db 2663 GCTGGCGCTGTGCTGCTCAATCCGACGCTCTATCTGACACCGTGTGCTGTGGCGCGC 2722

QY 507 GTGCGCGCGCAATACGCGGACACCGGACCGGTGGAATTTTTCGCGCTGGCGCGTTCCGCGC 566

Db 2723 CATCGCGCGCGCTTACGCGATGCGCGCAACGTTGGGCTTCCGCGCGCGCGCATGTGCGC 2782

QY 567 AAGCTGTGATCTGTTCCCGCTGTGGGTTCGGCGGCGAGCAGCATTTGTACGCCCGCTGTC 626

Db 2783 GTGATCTGTTCTGCTGCTGGGTTTCGGCGCACCGGCTGTGAGAGCGGTTGCTTGC 2842

QY 627 CAGCCCAAGTGTGGCGCTGATCAACGTCGTGGGCGAGTTGTGATGACCGCATTTGCG 686

Db 2843 CAGCCCGTTCGCTGGCGGCTGCTCGATGCGCTGATCGCGCGGCTGATGTGGCGCATCGC 2902

QY 687 CATCAACTGATGTTGATGGGTTAG 711

Db 2903 GCTGACGCTGCTGATGGCGCGGTTAG 2927

Search completed: March 15, 2004, 22:02:01
Job time : 2852.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:09:44 ; Search time 289.925 Seconds
(without alignments)
10418.122 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1016_1726

Perfect score: 711

Sequence: 1 atgggtgatcgaatctt.....aactgatgttgatgggttag 711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	711	100.0	711	8	ACC80941
2	711	100.0	822	4	AAF71779
3	711	100.0	822	4	AAS96098
4	711	100.0	2374	2	AAT96816
5	711	100.0	2374	9	ADB66196
6	711	100.0	34980	5	AAB68528
7	708	99.6	708	5	AAB68420
8	696.8	98.0	712	8	ACC80942
9	317.6	44.7	1568	4	AAH45375
10	60.4	8.5	636	3	AAH52691
11	57	8.0	15239	2	AAAT3356
12	57	8.0	110000	4	AAI99682_22
13	57	8.0	110000	4	AAI99683_22
14	56.2	7.9	86114	6	ABX09143
15	56.2	7.9	110000	4	AAI99682_05
16	56.2	7.9	110000	4	AAI99683_05
17	47.4	6.7	7521	5	AAS71378
18	47.4	6.7	7521	5	AAS94251
19	46.6	6.6	2000	7	ADA71938
20	44	6.2	2000	7	ADA71938
21	40	5.6	1173	7	ADA71370
22	39.4	5.5	624	7	ACF71727
23	39.4	5.5	110000	7	ACF67367_49

24	39.4	5.5	110000	7	ACF65387_0
25	39	5.5	993	4	AAF71777
26	39	5.5	993	4	AAS96096
27	39	5.5	2823	7	ACA38373
28	39	5.5	2826	7	ACA40558
29	39	5.5	110000	4	AAI99682_20
30	39	5.5	110000	4	AAI99683_20
31	38.6	5.4	1245	7	ACA26476
32	37.8	5.3	536	9	ADB68842
33	37.6	5.3	536	9	ADB68842
34	37.6	5.3	1149	7	ACF39391
35	37.6	5.3	110000	4	AAI99682_07
36	37.6	5.3	110000	4	AAI99683_07
37	37.4	5.3	1350	7	ACA27310
38	37.2	5.2	424	7	ABX47227
39	37.2	5.2	873	7	ACA37840
40	37.2	5.2	1509	7	ABX56070
41	37	5.2	945	7	ACA26592
42	37	5.2	1155	5	AAF26425
43	37	5.2	1464	7	ACA43878
44	36.8	5.2	678	7	ACA38495
45	36.8	5.2	681	7	ACA40313

ALIGNMENTS

RESULT 1					
ACC80941	ID	ACC80941 standard; DNA; 711 BP.			
XX	AC	ACC80941;			
XX	DT	27-OCT-2003 (revised)			
DT	11-AUG-2003	(first entry)			
XX	DE	Lyse protein encoding sequence.			
XX	KW	L-lysine; L-arginine; Lyse; ds.			
XX	OS	Corynebacterium glutamicum.			
XX	FH	Key Location/Qualifiers			
FT	CDS	1..711			
FT		/*tag= a			
FT		/product= "lyse protein"			
XX	PN	EP1266966-A2.			
XX	PD	18-DEC-2002.			
XX	PF	05-JUN-2002; 2002EP-00012539.			
XX	PR	12-JUN-2001; 2001JP-00177075.			
XX	PA	(AJIN) AJINOMOTO CO INC.			
XX	PI	Gunji Y, Yasueda H;			
XX	DR	WPI; 2003-241171/24.			
XX	DR	P-PSDB; ABR58213.			
XX	CC	Novel DNA encoding variant of Lyse protein from a coryneform bacterium, when introduced into methanol assimilating bacterium, facilitates excretion of L-lysine and/or L-arginine to outside of a cell.			
XX	PS	Example 1; Page 17-18; 23pp; English.			
XX	CC	The present invention relates to DNA encoding variants of protein with loop region and six hydrophobic helices which facilitates excretion of L-lysine and/or L-arginine to outside of cell of a methanol assimilating bacterium when introduced into the bacterium. The method is used for encoding a protein which facilitates excretion of L-lysine, L-arginine or			

CC both of these L-amino acids to outside of a cell of a methanol
CC assimilating bacterium when DNA of the method is introduced into the
CC bacterium. The present sequence represents a lyase protein from
CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003
CC to standardise OS field)
XX
SQ Sequence 711 BP; 135 A; 173 C; 222 G; 181 T; 0 U; 0 Other;

Query Match 100.0%; Score 711; DB 8; Length 711;
Best Local Similarity 100.0%; Pred. No. 3.9e-206;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCATCATCGAATCTTCATTACAGTCTGCTTTTGGGCGCAGCTTTTACTGTCC 60
Db |||||
QY 1 ATGTCATCATCGAATCTTCATTACAGTCTGCTTTTGGGCGCAGCTTTTACTGTCC 60
Db |||||
QY 61 ATCGGACCGCAGAGTACTGCTGATTAACAAGGAATTAAGCGCGAAGGACTCAATGCG 120
Db |||||
QY 121 GTTCTTCTGCTGTTTAATTTCTGACGCTTTTGTGTCATCGCGCACCTTGGGCGTT 180
Db |||||
QY 181 GATCTTTTGTCAATGCGCGCGATCGTCTCGATATTATGCTGGGCTGGCATCGCT 240
Db |||||
QY 181 GATCTTTTGTCAATGCGCGCGATCGTCTCGATATTATGCTGGGCTGGCATCGCT 240
Db |||||
QY 241 TACCTGTTATGCTTGGCGTATGCGAGCGAAAGCGCCATGACAAACAGGTGGAGCG 300
Db |||||
QY 241 TACCTGTTATGCTTGGCGTATGCGAGCGAAAGCGCCATGACAAACAGGTGGAGCG 300
Db |||||
QY 301 CCACAGATCATTTGAAGAACAGACACCGGTCGCCGATGACACGCTTTGGCGGTTGCG 360
Db |||||
QY 301 CCACAGATCATTTGAAGAACAGACACCGGTCGCCGATGACACGCTTTGGCGGTTGCG 360
Db |||||
QY 361 GCGGTGGCCACTGACACCGCGCAACCGGTCGCCGATGAGCGTGGATGACAGCAGCG 420
Db |||||
QY 361 GCGGTGGCCACTGACACCGCGCAACCGGTCGCCGATGAGCGTGGATGACAGCAGCG 420
Db |||||
QY 421 GTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTGAACCGAATCGGTAT 480
Db |||||
QY 421 GTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTGAACCGAATCGGTAT 480
Db |||||
QY 481 TTGACCGGTTGCTGTTTATCGCGCGTGGCGGCAATACCGCGACACCGCGTGG 540
Db |||||
QY 481 TTGACCGGTTGCTGTTTATCGCGCGTGGCGGCAATACCGCGACACCGCGTGG 540
Db |||||
QY 541 ATTTTCGCCGCTGGCGGCTTTCGCGCAAGCCTGATCTGTTCCCGTGGTGGTTTCGC 600
Db |||||
QY 541 ATTTTCGCCGCTGGCGGCTTTCGCGCAAGCCTGATCTGTTCCCGTGGTGGTTTCGC 600
Db |||||
QY 601 GCAGCAGATTGTCAGCGCGCTGTCAGCGCCCAAGTGGCGTGGATCAAGTGGTC 660
Db |||||
QY 601 GCAGCAGATTGTCAGCGCGCTGTCAGCGCCCAAGTGGCGTGGATCAAGTGGTC 660
Db |||||
QY 661 GTGGCAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGTTAG 711
Db |||||
QY 661 GTGGCAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGTTAG 711
Db |||||

RESULT 2

AAAF71779
ID AAF71779 standard; DNA; 822 BP.
XX
AC AAF71779;
XX
XX
DT 30-APR-2001 (first entry)
DE
DE Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:53.
XX
XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;

nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.

Corynebacterium glutamicum.

WO200100843-A2.

04-JAN-2001.

23-JUN-2000; 2000WO-IB000923.

25-JUN-1999; 99US-0141031P.

01-JUL-1999; 99DE-01030476.

02-JUL-1999; 99US-0142101P.

08-JUL-1999; 99DE-01031415.

08-JUL-1999; 99DE-01031418.

08-JUL-1999; 99DE-01031419.

08-JUL-1999; 99DE-01031420.

08-JUL-1999; 99DE-01031424.

08-JUL-1999; 99DE-01031428.

08-JUL-1999; 99DE-01031434.

08-JUL-1999; 99DE-01031435.

08-JUL-1999; 99DE-01031443.

08-JUL-1999; 99DE-01031453.

08-JUL-1999; 99DE-01031457.

08-JUL-1999; 99DE-01031465.

08-JUL-1999; 99DE-01031478.

08-JUL-1999; 99DE-01031510.

08-JUL-1999; 99DE-01031541.

08-JUL-1999; 99DE-01031573.

08-JUL-1999; 99DE-01031592.

08-JUL-1999; 99DE-01031632.

08-JUL-1999; 99DE-01031634.

08-JUL-1999; 99DE-01031636.

08-JUL-1999; 99DE-01032125.

09-JUL-1999; 99DE-01032136.

09-JUL-1999; 99DE-01032130.

09-JUL-1999; 99DE-01032186.

09-JUL-1999; 99DE-01032206.

09-JUL-1999; 99DE-01032227.

09-JUL-1999; 99DE-01032228.

09-JUL-1999; 99DE-01032229.

09-JUL-1999; 99DE-01032230.

14-JUL-1999; 99DE-01032922.

14-JUL-1999; 99DE-01032926.

14-JUL-1999; 99DE-01032928.

14-JUL-1999; 99DE-01033004.

14-JUL-1999; 99DE-01033005.

14-JUL-1999; 99DE-01033006.

14-JUL-1999; 99US-0148613P.

27-AUG-1999; 99DE-01040764.

27-AUG-1999; 99DE-01040765.

27-AUG-1999; 99DE-01040766.

27-AUG-1999; 99DE-01040832.

31-AUG-1999; 99DE-01041378.

31-AUG-1999; 99DE-01041379.

31-AUG-1999; 99DE-01041380.

31-AUG-1999; 99DE-01041394.

31-AUG-1999; 99DE-01041396.

03-SEP-1999; 99DE-01042076.

03-SEP-1999; 99DE-01042077.

03-SEP-1999; 99DE-01042079.

03-SEP-1999; 99DE-01042086.

03-SEP-1999; 99DE-01042087.

03-SEP-1999; 99DE-01042088.

03-SEP-1999; 99DE-01042095.

03-SEP-1999; 99DE-01042124.

03-SEP-1999; 99DE-01042129.

09-MAR-2000; 2000US-0187970P.

(BADI) BASF AG.

XX

XX

PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 XX WPI: 2001-137957/14.
 DR P-PSDB; AAB79660.
 XX
 PT Nucleic acids from *Corynebacterium glutamicum* encoding metabolic pathway
 PT proteins, useful for producing fine chemicals in microorganisms,
 PT including organic acids, nonproteinogenic amino acids, and purine and
 PT pyrimidine bases.
 XX
 PS Claim 3; Page 233-234; 1737pp; English.
 XX
 CC AAF71753 to AAF72330 encode the *Corynebacterium glutamicum* metabolic
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
 CC nucleic acids are useful for the production of fine chemicals in
 CC microorganisms, including organic acids, nonproteinogenic amino acids,
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
 CC and unsaturated fatty acids, diols, carboxylates, aromatic compounds,
 CC vitamins, cofactors, polyketides and enzymes
 XX
 SQ Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;
 Query Match 100.0%; Score 711; DB 4; Length 822;
 Best Local Similarity 100.0%; Pred. No. 4.2e-206;
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGTGATCATGGAATCTTCTATTACAGGTCTGTTTGGGGCCAGTCTTTTACTGTCC 60
 DB 92 ATGGTGATCATGGAATCTTCTATTACAGGTCTGTTTGGGGCCAGTCTTTTACTGTCC 151
 QY 61 ATCGACCGCAGAGATGATCTGGTGAATTAACAAGAAATTAAGCGGAGACATCATTCGG 120
 DB 152 ATCGACCGCAGAGATGATCTGGTGAATTAACAAGAAATTAAGCGGAGACATCATTCGG 211
 QY 121 GTTCTCTCGTGTGTTTAATTTCTGACGTCTTTTCTCATCGCGGCGACCTTTGGCGTT 180
 DB 212 GTTCTCTCGTGTGTTTAATTTCTGACGTCTTTTCTCATCGCGGCGACCTTTGGCGTT 271
 QY 181 GATCTTTTGTCCAAATGCGCGCGATCGTGTGATTAATATGCGTGGGGTGGCATCGCT 240
 DB 272 GATCTTTTGTCCAAATGCGCGCGATCGTGTGATTAATATGCGTGGGGTGGCATCGCT 331
 QY 241 TACCTGTATGGTTGGCGTCAATGCGACGAAAGACCATGACAAACAGGTGGAAGCG 300
 DB 332 TACCTGTATGGTTGGCGTCAATGCGACGAAAGACCATGACAAACAGGTGGAAGCG 391
 QY 301 CCACAGATCATTGAAGAAACAGACCAACCGTCCGATGACAGCTTTGGGCGGTTCCG 360
 DB 392 CCACAGATCATTGAAGAAACAGACCAACCGTCCGATGACAGCTTTGGGCGGTTCCG 451
 QY 361 GCGGTGGCCACTGACACGCGCAACCGGCGTGGGTTGAGGTGAGCGTCGATAGCAGCGG 420
 DB 452 GCGGTGGCCACTGACACGCGCAACCGGCGTGGGTTGAGGTGAGCGTCGATAGCAGCGG 511
 QY 421 GTTTGGTGAAGCCCATTTGATGCAATGCTGCTGCTGCTGATTAATATGCGTGGGGTGGCATCGCT 480
 DB 512 GTTTGGTGAAGCCCATTTGATGCAATGCTGCTGCTGCTGATTAATATGCGTGGGGTGGCATCGCT 571
 QY 481 TTGACGCGGTTTGTGTTTATCGCGCGCTGCGGCGCAATACGCGCACCGGACGCGTGG 540
 DB 572 TTGACGCGGTTTGTGTTTATCGCGCGCTGCGGCGCAATACGCGCACCGGACGCGTGG 631
 QY 541 ATTTTTCGCGCTGGCGGTTTCGCGCAAGCCTGATCTGGTTCCCGCTGGTGGGTTTCGCGC 600
 DB 632 ATTTTTCGCGCTGGCGGTTTCGCGCAAGCCTGATCTGGTTCCCGCTGGTGGGTTTCGCGC 691
 QY 601 GCAGACGATTGTCAACCGCGCTGTGCGAGCCCAAGGTTGGCGCTGATCAAGTCGTC 660
 DB 692 GCAGACGATTGTCAACCGCGCTGTGCGAGCCCAAGGTTGGCGCTGATCAAGTCGTC 751
 QY 661 GTGCGAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 711
 DB 752 GTGCGAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 802

RESULT 3
 AAS96098
 ID AAS96098 standard; DNA; 822 BP.
 XX
 AC AAS96098;
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE C. glutamicum gene #23 encoding metabolic pathway protein.
 XX
 KW Metabolic pathway protein; MP; lysine biosynthesis pathway;
 KW methionine biosynthesis pathway; large-scale production of fine chemical;
 KW *Corynebacterium diphtheriae*; diphtheria; ds.
 XX
 OS *Corynebacterium glutamicum*.
 XX
 PN WO200166573-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 22-DEC-2000; 2000WO-IB002035.
 XX
 PR 09-MAR-2000; 2000US-0187970P.
 PR 23-JUN-2000; 2000US-00606740.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 PI Kim J, Lee H, Hwang B;
 XX
 DR WPI: 2001-582269/65.
 DR P-PSDB; AAU71888.
 XX
 PT Nucleic acids encoding metabolic pathway proteins from *Corynebacterium*
 PT *glutamicum*, useful for producing methionine and lysine in *Corynebacterium*
 PT and *Brevibacterium*.
 PT
 PS Disclosure; Page 214-215; 316pp; English.
 XX
 CC The present invention relates to the isolation of novel *Corynebacterium*
 CC *glutamicum* genes encoding metabolic pathway (MP) proteins (AAU71863-
 CC AAU71922). The metabolic pathway proteins of the invention include
 CC enzymes involved in the lysine and methionine biosynthetic pathways. The
 CC polynucleotide sequences of the invention can be used for the large-scale
 CC production and/or modulation of expression of fine chemicals such as
 CC lysine and methionine. The sequences of the invention may be used to
 CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a
 CC subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum
 CC genes encoding the novel metabolic pathway proteins of the invention.
 XX
 SQ Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;
 Query Match 100.0%; Score 711; DB 4; Length 822;
 Best Local Similarity 100.0%; Pred. No. 4.2e-206;
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGTGATCATGGAATCTTCTATTACAGGTCTGTTTGGGGCCAGTCTTTTACTGTCC 60
 DB 92 ATGGTGATCATGGAATCTTCTATTACAGGTCTGTTTGGGGCCAGTCTTTTACTGTCC 151
 QY 61 ATCGACCGCAGAGATGATCTGGTGAATTAACAAGAAATTAAGCGGAGACATCATTCGG 120
 DB 152 ATCGACCGCAGAGATGATCTGGTGAATTAACAAGAAATTAAGCGGAGACATCATTCGG 211
 QY 121 GTTCTCTCGTGTGTTTAATTTCTGACGTCTTTTCTCATCGCGGCGACCTTTGGCGTT 180
 DB 212 GTTCTCTCGTGTGTTTAATTTCTGACGTCTTTTCTCATCGCGGCGACCTTTGGCGTT 271
 QY 181 GATCTTTTGTCCAAATGCGCGCGATCGTGTGATTAATATGCGTGGGGTGGCATCGCT 240
 DB 272 GATCTTTTGTCCAAATGCGCGCGATCGTGTGATTAATATGCGTGGGGTGGCATCGCT 331

QY 241 TACCTGTTATGTTTGGCGTCATGCGAGCAAGAACCCCATGACAAACAAAGGTGGAAGCG 300
Db |||||
QY 332 TACCTGTTATGTTTGGCGTCATGCGAGCAAGAACCCCATGACAAACAAAGGTGGAAGCG 391
Db |||||
QY 301 CCACAGATCATTTGAAGAACAGAACCAACCGTGCCTGATGACACGCGTTCGGGCGTTGCG 360
Db |||||
QY 332 CCACAGATCATTTGAAGAACAGAACCAACCGTGCCTGATGACACGCGTTCGGGCGTTGCG 451
Db |||||
QY 361 GCGGTGGCCACTGACACGCGCAACCGGTTGCGGTTGAGGTTGAGGTCGATGAAGCAGCGG 420
Db |||||
QY 452 GCGGTGGCCACTGACACGCGCAACCGGTTGCGGTTGAGGTTGAGGTCGATGAAGCAGCGG 511
Db |||||
QY 421 GTTTGGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGTTGAACCCCAATGCGTAT 480
Db |||||
QY 512 GTTTGGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGTTGAACCCCAATGCGTAT 571
Db |||||
QY 481 TTGGAGCGGTTTGTGTTATCGCGCGCTGCGCGCAATACGCGCAGACACCGGACGCTGG 540
Db |||||
QY 572 TTGGAGCGGTTTGTGTTATCGCGCGCTGCGCGCAATACGCGCAGACACCGGACGCTGG 631
Db |||||
QY 541 ATTTTTCGCGCTGCGCGCTTTCGCGCAAGCCTGATCTGTTCCCGCTGCTGGTTTCGCG 600
Db |||||
QY 632 ATTTTTCGCGCTGCGCGCTTTCGCGCAAGCCTGATCTGTTCCCGCTGCTGGTTTCGCG 691
Db |||||
QY 601 GCACGACATTTGACGCGCGCTGTCAGCGCCAGGTTGCGGCTGGATCAAGTCGTC 660
Db |||||
QY 692 GCACGACATTTGACGCGCGCTGTCAGCGCCAGGTTGCGGCTGGATCAAGTCGTC 751
Db |||||
QY 661 GTGCGCATTTGATGACGCGCAATTTGGCCATCAAACTGATTTGATGGGTTAG 711
Db |||||
QY 752 GTGCGCATTTGATGACGCGCAATTTGGCCATCAAACTGATTTGATGGGTTAG 802
Db |||||
RESULT 4
AAT96816
ID AAT96816 standard; DNA; 2374 BP.
XX AAT96816;
XX
XX 12-MAR-1998 (first entry)
XX
XX DNA encoding LysG, LysE and ORF3 from Corynebacterium glutamicum.
XX
XX LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;
KW Microbial production; amino acid; animal feed additive; ds.
XX
XX Corynebacterium glutamicum.
FH Key Location/Qualifiers
CDS complement(82..954)
FT /*tag= a
FT /label= LysG
FT 1016..1726
FT /*tag= b
FT /label= LysE
FT complement(1723..2373)
FT /*tag= c
FT /label= orf3
XX
XX DE19548222-A1.
XX
XX 26-JUN-1997.
XX
XX 22-DEC-1995; 95DE-01048222.
XX
XX 22-DEC-1995; 95DE-01048222.
XX (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
XX Vrijic M, Eggeling L, Sahm H;
XX WPI; 1997-333867/31.
XX
XX

DR P-PSDB; AAW37714, AAW37715, AAW37716.
XX
XX Increasing microbial production of amino acids, especially lysine - by
PT improving export carrier activity or corresponding gene expression, also
PT new export and regulatory genes from Corynebacterium.
XX
XX Claim 23 and 26; Page; 16pp; German.
XX
XX This DNA, isolated from Corynebacterium glutamicum, contains the LysG,
CC LysE and ORF3 genes. LysG and LysE encode a lysine transport regulatory
CC protein and an export protein, respectively. Microbial production of
CC amino acids (A) is improved by increasing the export-carrier activity
CC and/or the export gene expression in a microorganism that produces (A).
CC The method is specifically used to increase production of lysine, used as
CC an animal feed additive. Other (A) are variously useful as
CC pharmaceuticals, condiments and intermediates for fine chemicals. This
CC method increases the amount of (A) secreted into the culture medium.
CC Export of (A) has been found to depend on a single gene. NB. This
CC sequence has been created from the information given in table 2 of the
CC specification
XX
SQ Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;
Query Match 100.0%; Score 711; DB 2; Length 2374;
Best Local Similarity 100.0%; Pred. No. 6.9e-206;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTGATCATGGAATCTTCATACAGGTCTCTTTTGGGGCCAGTCTTTTACTGTCC 60
Db 1016 ATGCTGATCATGGAATCTTCATACAGGTCTCTTTTGGGGCCAGTCTTTTACTGTCC 1075
QY 61 ATCGGACCGCAGAAATGTAATGATTAACAAAGGAATTAAGCGCGAAGACTCATTGCG 120
Db 1076 ATCGGACCGCAGAAATGTAATGATTAACAAAGGAATTAAGCGCGAAGACTCATTGCG 1135
QY 121 GTTCTTCTCGTGTGTTTAAATTTCTGAGCTCTTTTGTTCATCGCGCCACCTTGGCGTT 180
Db 1136 GTTCTTCTCGTGTGTTTAAATTTCTGAGCTCTTTTGTTCATCGCGCCACCTTGGCGTT 1195
QY 181 GATCTTTTGTCCAATGCGCGCGCATCTGCTCGATATTATGCTGCGTGGGCTGGCATCGCT 240
Db 1196 GATCTTTTGTCCAATGCGCGCGCATCTGCTCGATATTATGCTGCGTGGGCTGGCATCGCT 1255
QY 241 TACCTGTTATGTTTCCCGTCATGCGCAGCAAGAGCGCCATGACAAACAAAGGTGGAAGCG 300
Db 1256 TACCTGTTATGTTTCCCGTCATGCGCAGCAAGAGCGCCATGACAAACAAAGGTGGAAGCG 1315
QY 301 CCACAGATCATTTGAAGAACAGAACCAACCGTGCCTGATGACACGCGTTCGGGCGTTGCG 360
Db 1316 CCACAGATCATTTGAAGAACAGAACCAACCGTGCCTGATGACACGCGTTCGGGCGTTGCG 1375
QY 361 GCGGTGGCCACTGACACGCGCAACCGGTTGCGGTTGAGGTTGAGGTCGATGAAGCAGCGG 420
Db 1376 GCGGTGGCCACTGACACGCGCAACCGGTTGCGGTTGAGGTTGAGGTCGATGAAGCAGCGG 1435
QY 421 GTTTGGGTAAAGCCCATGTTGATGCGCAATCGTCTGATGTTGAACCCCAATGCGTAT 480
Db 1436 GTTTGGGTAAAGCCCATGTTGATGCGCAATCGTCTGATGTTGAACCCCAATGCGTAT 1495
QY 481 TTGGAGCGGTTTGTGTTATCGCGCGCTGCGCGCAATACGCGCAGACACCGGACGCTGG 540
Db 1496 TTGGAGCGGTTTGTGTTATCGCGCGCTGCGCGCAATACGCGCAGACACCGGACGCTGG 1555
QY 541 ATTTTTCGCGCTGCGCGCTTTCGCGCAAGCCTGATCTGTTCCCGCTGCTGGGTTTCGCG 600
Db 1556 ATTTTTCGCGCTGCGCGCTTTCGCGCAAGCCTGATCTGTTCCCGCTGCTGGGTTTCGCG 1615
QY 601 GCACGACATTTGTCAGCGCGCTGTCAGCGCCCAAGGTTGCGGCTGGATCAACGCTGTC 660
Db 1616 GCACGACATTTGTCAGCGCGCTGTCAGCGCCCAAGGTTGCGGCTGGATCAACGCTGTC 1675
QY 661 GTGCGCATTTGATGACGCGCAATTTGGCCATCAAACTGATTTGATGGGTTAG 711
Db |||||

XX PS Disclosure; SEQ ID NO 7063; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein

XX CC sequences from the Corynebacterium glutamicum. These

XX CC are useful for identifying the mutation point of a gene derived from a

XX CC mutant of corynebacterium, measuring expression amount and analysing

XX CC the expression profile or expression pattern of a gene derived from

XX CC Corynebacterium, and identifying a homologue of a gene derived from

XX CC corynebacterium. Corynebacterium are useful for producing amino

XX CC acids, nucleic acids, vitamins, saccharides and organic acids,

XX CC particularly L-lysine. The present sequence is a nucleic acid described

XX CC in the exemplification of the invention. Note: The sequence data for this

XX CC patent did not form part of the printed specification, but was obtained

XX CC in electronic format directly from the European Patent Office

XX SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;

Query Match 100.0%; Score 711; DB 5; Length 349980;

Best Local Similarity 100.0%; Pred. No. 7.2e-205; Indels 0; Gaps 0;

Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGATCATGGAATCTTCATTACAGCTCTGCTTTTGGGGCCAGCTTTTACTGTCC 60

Db 128953 ATGGTGATCATGGAATCTTCATTACAGCTCTGCTTTTGGGGCCAGCTTTTACTGTCC 128894

QY 61 ATCGGACCGCAGATGACTGTGATTAACAGGAATTAAGCGGAGGACTCATTCGG 120

Db 128893 ATCGGACCGCAGATGACTGTGATTAACAGGAATTAAGCGGAGGACTCATTCGG 128834

QY 121 GTTCTCTCGTGTTAAATTTCTGACGCTCTTTTTCATCCCGGACCTTGGGCGTT 180

Db 128833 GTTCTCTCGTGTTAAATTTCTGACGCTCTTTTTCATCCCGGACCTTGGGCGTT 128774

QY 181 GATCTTTTGTCAATGCCGCGCATCGTCTCGATATTATGCGCTGCGGTGGCATCGT 240

Db 128773 GATCTTTTGTCAATGCCGCGCATCGTCTCGATATTATGCGCTGCGGTGGCATCGT 128714

QY 241 TACTGTATGTTTGGCGTCATGCGGAGGAGCCGATGACAAAGCTGGAAGCG 300

Db 128713 TACTGTATGTTTGGCGTCATGCGGAGGAGCCGATGACAAAGCTGGAAGCG 128654

QY 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGCCCGATGACACGCTTTGGGCGTTG 360

Db 128653 CCACAGATCATTTGAAGAAACAGAACCAACCGTGCCCGATGACACGCTTTGGGCGTTG 128594

QY 361 GCGGTGGCCACTGACACGCGCAACCGGCTGCGGTTGGAGTGAAGCTGATAGCAGCG 420

Db 128593 GCGGTGGCCACTGACACGCGCAACCGGCTGCGGTTGGAGTGAAGCTGATAGCAGCG 128534

QY 421 GTTTGGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGTTGAACCCGATGCGTAT 480

Db 128533 GTTTGGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGTTGAACCCGATGCGTAT 128474

QY 481 TTGACACGCTTTGTGTTATCGGGCGCTGCGCGGCAATACGGCGACACCGGAGTGG 540

Db 128473 TTGACACGCTTTGTGTTATCGGGCGCTGCGCGGCAATACGGCGACACCGGAGTGG 128414

QY 541 ATTTTCGCGCTGGCGCTGCGCGGCAAGCTGATCTGGTTCGCTGGTGGTTCGCG 600

Db 128413 ATTTTCGCGCTGGCGCTGCGCGGCAAGCTGATCTGGTTCGCTGGTGGTTCGCG 128354

QY 601 GCACGACGATTTGACGCGCTGTCGACGCGCCAGGCTGCGCTGGATCAACGTCGTC 660

Db 128353 GCACGACGATTTGACGCGCTGTCGACGCGCCAGGCTGCGCTGGATCAACGTCGTC 128294

QY 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 711

Db 128293 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 128243

ID XX AAH68420 standard; DNA; 708 BP.

XX AC AAH68420;

XX DT 26-SEP-2001 (first entry)

XX DE C glutamicum coding sequence fragment SEQ ID NO: 3455.

XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;

XX KW organic acid synthesis; ds.

XX OS Corynebacterium glutamicum.

XX PN BP108790-A2.

XX PX 20-JUN-2001.

XX PF 19-DEC-2000; 2000EP-00127688.

XX PR 16-DEC-1999; 99JP-00377484.

XX PR 07-APR-2000; 2000JP-00159162.

XX PR 03-AUG-2000; 2000JP-00280988.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX PX WPI; 2001-376931/40.

XX DR P-PSDB; AAG93201.

XX PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying

XX PT mutation point of a gene, measuring expression of a gene, analyzing

XX PT expression profile or pattern of a gene and identifying homologous gene.

XX PS Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein

XX CC sequences from the Corynebacterium glutamicum. These

XX CC are useful for identifying the mutation point of a gene derived from a

XX CC mutant of corynebacterium, measuring expression amount and analysing

XX CC the expression profile or expression pattern of a gene derived from

XX CC Corynebacterium, and identifying a homologue of a gene derived from

XX CC corynebacterium. Corynebacterium are useful for producing amino

XX CC acids, nucleic acids, vitamins, saccharides and organic acids,

XX CC particularly L-lysine. The present sequence is a nucleic acid described

XX CC in the exemplification of the invention. Note: The sequence data for this

XX CC patent did not form part of the printed specification, but was obtained

XX CC in electronic format directly from the European Patent Office

XX SQ Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;

Query Match 99.6%; Score 708; DB 5; Length 708;

Best Local Similarity 100.0%; Pred. No. 3.2e-205; Indels 0; Gaps 0;

Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGATCATGGAATCTTCATTACAGCTCTGCTTTTGGGGCCAGCTTTTACTGTCC 60

Db 1 ATGGTGATCATGGAATCTTCATTACAGCTCTGCTTTTGGGGCCAGCTTTTACTGTCC 60

QY 61 ATCGGACCGCAGATGACTGTGATTAACAGGAATTAAGCGGAGGACTCATTCGG 120

Db 61 ATCGGACCGCAGATGACTGTGATTAACAGGAATTAAGCGGAGGACTCATTCGG 120

QY 121 GTTCTCTCGTGTTAAATTTCTGACGCTCTTTTTCATCCCGGACCTTGGGCGTT 180

Db 121 GTTCTCTCGTGTTAAATTTCTGACGCTCTTTTTCATCCCGGACCTTGGGCGTT 180

QY 181 GATCTTTTGTCAATGCCGCGCATCGTCTCGATATTATGCGCTGCGGTGGCATCGT 240

Db 181 GATCTTTTGTCAATGCCGCGCATCGTCTCGATATTATGCGCTGCGGTGGCATCGT 240

QY 241 TACTGTATGTTTGGCGTCATGCGGAGGAGCCGATGACAAAGCTGGAAGCG 300

Db 241 TACCTGATGTTGCGCTCATGCGAAGAGCGCCATGACAAACAGGTGGAAGCG 300
Qy 301 CCACAGATCATTTGAAAGAACAGAACACCGTGGCCGATGACACGCTTTTGGCGGTTCG 360
Db 301 CCACAGATCATTTGAAAGAACAGAACACCGTGGCCGATGACACGCTTTTGGCGGTTCG 360
Qy 361 GCGGTGGCCACTGACACGCGCAACCGGTGGCGGTGGAGGTGACGCTTCATAGCGGG 420
Db 361 GCGGTGGCCACTGACACGCGCAACCGGTGGCGGTGGAGGTGACGCTTCATAGCGGG 420
Qy 421 GTTTCGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGGTTGAACCCGATTCGTTAT 480
Db 421 GTTTCGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGGTTGAACCCGATTCGTTAT 480
Qy 481 TTGGACGCGTTTGTGTTTATCGCGCGCTCGCGCGCAATACGCGCACACCGGACGTTGG 540
Db 481 TTGGACGCGTTTGTGTTTATCGCGCGCTCGCGCGCAATACGCGCACACCGGACGTTGG 540
Qy 541 ATTTCGCGCTGGCGGTTCGCGGCAAGCTGATCTGGTTCCCGCTGGTGGTTCGGC 600
Db 541 ATTTCGCGCTGGCGGTTCGCGGCAAGCTGATCTGGTTCCCGCTGGTGGTTCGGC 600
Qy 601 GCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGCAAGGTGTGGCGCTGGATCAACGTGTC 660
Db 601 GCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGCAAGGTGTGGCGCTGGATCAACGTGTC 660
Qy 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGT 708
Db 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGT 708

RESULT 8

ACC80942
ID ACC80942 standard; DNA; 712 BP.
XX
AC ACC80942;
XX
DT 27-OCT-2003 (revised)
DT 11-AUG-2003 (first entry)
XX
LysE24 protein encoding sequence.
XX
L-lysine, L-arginine; LysE24; da.
XX
Corynebacterium glutamicum.
XX
Key Location/Qualifiers
CDS 1..711
FT /*tag= a
FT /product= "lysE protein"
FT
XX
EP1266966-A2.
XX
18-DEC-2002.
XX
05-JUN-2002; 2002EP-00012539.
XX
12-JUN-2001; 2001JP-00177075.
XX
(AJIN) AJINOMOTO CO INC.
XX
Gunji Y, Yasueda H;
XX
WPI; 2003-241171/24.
DR P-PSDB; ABR58214.
XX
Novel DNA encoding variant of LysE protein from a coryneform bacterium,
PT when introduced into methanol assimilating bacterium, facilitates
XX excretion of L-lysine and/or L-arginine to outside of a cell.
XX
Example 1; Page 19-20; 23pp; English.
XX

CC The present invention relates to DNA encoding variants of protein with
CC loop region and six hydrophobic helices which facilitates excretion of L-
CC lysine and/or L-arginine to outside of cell of a methanol assimilating
CC bacterium when introduced into the bacterium. The method is used for
CC encoding a protein which facilitates excretion of L-lysine, L-arginine or
CC both of these L-amino acids to outside of a cell of a methanol
CC assimilating bacterium when DNA of the method is introduced into the
CC bacterium. The present sequence represents a LysE24 protein from
CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003
CC to standardise OS field)
XX

SQ Sequence 712 BP; 133 A; 173 C; 224 G; 182 T; 0 U; 0 Other;

Query Match 98.0%; Score 696.8; DB 8; Length 712;
Best Local Similarity 99.6%; Pred. No. 8.3e-202;
Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 ATGGTGATCATGGAATCTTCATTACAGGTCTGCTTTGGGGGCGAGCTTTTACGTGCC 60
Db 1 ATGGTGATCATGGAATCTTCATTACAGGTCTGCTTTGGGGGCGAGCTTTTGTCTGCC 60
Qy 61 ATCGACCGCAGATGATCTGCTGATTAACAGGAATTAAGCGCGAGGACTCATTTGG 120
Db 61 ATCGACCGCAGATGATCTGCTGATTAACAGGAATTAAGCGCGAGGACTCATTTGG 120
Qy 121 GTTCTTCTGCTGTTTAAATTTCTGACGCTCTTTTCTCATCGCGGCACTTTGGCGTT 180
Db 121 GTTCTTCTGCTGTTTAAATTTCTGACGCTCTTTTCTCATCGCGGCACTTTGGCGTT 180
Qy 181 GATCTTTTGTCCAAATGCGCGCGGATCTGCTGCTGATTAATGCGTGGGTGGCATCGT 240
Db 181 GATCTTTTGTCCAAATGCGCGCGGATCTGCTGCTGATTAATGCGTGGGTGGCATCGT 240
Qy 241 TACCTGTTATGTTTGGCTGATGCGAGGAGAGCCATGACAAACAGGTGGAAGCG 300
Db 241 TACCTGTTATGTTTGGCTGATGCGAGGAGAGAGCCATGACAAACAGGTGGAAGCG 300
Qy 301 CCACAGATCATTTGAAGAACAGAACCAACCGTGGCCGATGACACGCTTTTGGCGGT 359
Db 301 CCACAGATCATTTGAAGAACAGAACCAACCGTGGCCGATGACACGCTTTTGGCGGT 360
Qy 360 GCGGTGGCCACTGACACGCGCAACCGGTGGCGGTGGAGGTGAGCTGCGATAGCAGCG 419
Db 361 GCGGTGGCCACTGACACGCGCAACCGGTGGCGGTGGAGGTGAGCTGCGATAGCAGCG 420
Qy 420 GGTTCGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGTTGAACCCGATTCGTA 479
Db 421 GGTTCGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGTTGAACCCGATTCGTA 480
Qy 480 TTTGGACGCGTTTGTGTTTATCGCGCGCTCGCGCGCAATACGCGGACACCGGACGTTG 539
Db 481 TTTGGACGCGTTTGTGTTTATCGCGCGCTCGCGCGCAATACGCGGACACCGGACGTTG 540
Qy 540 GATTTTCGCGCTGGCGGTTCGCGGCAAGCTGATCTGCTGCTGGCTGGGTTCGG 599
Db 541 GATTTTCGCGCTGGCGGTTCGCGGCAAGCTGATCTGCTGCTGGCTGGGTTCGG 600
Qy 600 CGCAGCAGCATTTGTCACGCGCGCTGTCAGCGCCCAAGGCTGGCGCTGCATCAACGTCGT 659
Db 601 CGCAGCAGCATTTGTCACGCGCGCTGTCAGCGCCCAAGGCTGGCGCTGCATCAACGTCGT 660
Qy 660 CGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 711
Db 661 CGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 712

RESULT 9

AAH45375
ID AAH45375 standard; DNA; 1568 BP.
XX
AC AAH45375;
XX
DT 11-SEP-2001 (first entry)
XX

XX C. thermoaminogenes lysin biosynthetic enzyme lyse DNA.
XX Heat-resistant; lysin biosynthesis; enzyme: coryneform;
KW aspartate-semialdehyde dehydrogenase; lyse; ds.
XX Corynebacterium thermoaminogenes.
XX JP2001120270-A.
XX 08-MAY-2001.
XX 01-NOV-1999; 99JP-003111148.
XX 01-NOV-1999; 99JP-003111148.
XX (AJIN) AJINOMOTO KK.
XX WPI; 2001-364760/38.
XX P-PSDB; RAG64047.
XX A heat-resistant lysin biosynthetic system enzyme gene of a high
PT temperature-resistant coryneform microbe.
XX Example 5; Page 22-24; 27pp; Japanese.
XX The invention relates to a gene from a high temperature-resistant
CC coryneform microbe that encodes a heat-resistant lysin biosynthetic
CC enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and
CC can be used for growing amino acid-producing microbes. The present
CC sequence encodes an enzyme of the invention
XX
XX Sequence 1568 BP; 258 A; 525 C; 490 G; 295 T; 0 U; 0 Other;
Query Match 44.7%; Score 317.6; DB 4; Length 1568;
Best Local Similarity 67.7%; Pred. No. 4.8e-86;
Matches 481; Conservative 0; Mismatches 214; Indels 15; Gaps 2;
QY 1 ATGTGATATGGAATCTTACATGAGTCTGCTTTGGGGCCAGTCTTTTACTGTCC 60
Db 778 ATGGGACATGGAATCTTGTGACCGGTTGTTGGAGCCAGTCTGTTGGCC 837
QY 61 ATCGACCGCAGAGTACTGCTGATTAACAGAGATTAAGCGCGAGGACTATTGCG 120
Db 838 ATCGGCCACAGAGTCTGCTGATCAACAGGCGCATCAACGCGAGGCGATCACGCC 897
QY 121 GTTCTTCTGCTGTTTAACTTCTGAGTCTTTTGTTCATCGCCGCACTTGGCGGT 180
Db 898 GTCATATCTGCTGCTGCTGCGAGTGTCTGCTTCCCTCGCACTCTGCGGATCGCC 957
QY 181 GATCTTTTGTCCAAATGCGCGCGATCGTCTCGATATTATGCGTGGGGTGGCATCGT 240
Db 958 GGCCTGATCTCCGACACCGCCCGATCATTTCTCGACATCTCTGCTGTGGCGCATGCC 1017
QY 241 TACCTGTATGTTTCCGTCATGCGAGCAAGAGCGCATGACAAACAGGTGAGCG 300
Db 1018 TACCTGTGTGTTTCCGTCATGATGCGCGCGCGCGCGCTGCGCGCCCGCACGAGGTA 1077
QY 301 CCACAGATCATTAAGAAACAGAACCAACCGTCCCGATGACAGCGCTTTGGGGGTTG 360
Db 1078 ACCTTTGTC-----GAGCATTCGAACCGTTG---CGGACGGTCCGCTCC 1122
QY 361 GCGGTGGCCACTGACACGCGCAACCGGTCGGGTGGAGGTGAGCGTGCATTAAGCAGGG 420
Db 1123 GCGGGGGGCTGACGACGAGAAACACGACCCCGGCTCCGATCATCAGGACCCCGGCG 1182
QY 421 GTTGGGTAAAGCCATGTTGATGGCAATCGTCTGACCTGTTGACCGGATCGGTAT 480
Db 1183 GTCTGGGTCAAGCCCATGCTCATGGCCATTGTGCTGACCTCAATCCCAATGCCATAC 1242
QY 481 TTGAGCGGTTGTTGTTTATCGCGCGCTCGGCGCAATACGCGACACCGGAGTGG 540
Db 1243 CTGGATGCTTCTGCTTCTATCGTGTGTTGAGGCCAGTACGCGGAGACCGGTGG 1302

QY 541 ATTTTCGCGCTGCGGCTTCGGGGAAGCCTGATCTGCTTCCGCTGCTGGGTTTCGSC 600
Db 1303 ATCTTGGTGGGGTGGCTTCGCGCCAGCCTGCTCTGGTTCCCACTGCTGGTTACGGC 1362
QY 601 GCAGCAGCATTTGTACCGCGCTGTCAGCCCCCAAGGTGTGGCGTGGATCAAGCTGTC 660
Db 1363 GCGGCCGCACTGTCTGCTGCTCCCTGTCAGCCCGCGGCTCTGGCGCTGATCAATAGT 1422
QY 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGTTA 710
Db 1423 GTGGCGGTGTCTCAGCGGATTTGGCGGTGAAGCTGATCTGATGGTTA 1472
RESULT 10
AAA52691
ID AAA52691 standard; DNA; 636 BP.
XX AC AAA52691;
XX DT 03-JAN-2001 (first entry)
XX DE Escherichia coli yggA gene.
XX KW E. coli; yggA gene; amino acid production; excretion protein gene;
XX amino acid excretion protein; ds.
XX OS Escherichia coli.
XX FH Key Location/Qualifiers
XX FT 1. .636
XX FT CDS /tag= a
XX FT /product= "yggA"
XX PN EP1016710-A2.
XX PD 05-JUL-2000.
XX PF 17-DEC-1999; 99EP-00125263.
XX PR 30-DEC-1999; 98RU-00124016.
XX PR 09-MAR-1999; 99RU-00104431.
XX RA (AJIN) AJINOMOTO CO INC.
XX PI Livshits VA, Zakataeva NP, Nakanishi K, Aleshin VV, Troshin PV;
PI Tokhmakova IL;
XX WPI; 2000-414802/36.
XX P-PSDB; AAB01789.
XX Increased production of L-amino acids by an Escherichia bacterium
PT comprises increasing the expression amount of an L-amino acid excretion
PT protein.
XX PS Disclosure; Page 24; 29pp; English.
XX CC The present sequence is the yggA gene (an excretion protein gene) of
CC Escherichia coli. The amino acid excretion protein produced from this
CC gene is involved in the production of amino acids, and an increase in its
CC expression leads to an increased accumulation of amino acids in the cell.
CC In this case, an increase in arginine, glutamic acid and lysine is
CC achieved if multiple copies of the gene are transfected into a bacterium.
CC The bacterium used is E. coli
XX
XX Sequence 636 BP; 120 A; 143 C; 180 G; 193 T; 0 U; 0 Other;
Query Match 8.5%; Score 60.4; DB 3; Length 636;
Best Local Similarity 51.1%; Pred. No. 8.2e-08;
Matches 142; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
QY 19 TTCTATACAGTCTGCTTTTGGGCGCCAGTCTTTTACTCTCATCGACCGCATGTA 78

Db 13 TACTTTCAAGCTCTTGGCTGCGGCGCTATGCTACCGCTCGCTCCACAAAATGCT 72
 Qy 79 CTGGTGATTAACAAGGATTAAGCGGAGGACTCATTTGGCTCTTCTGCTGTGTTTA 138
 Db 73 TTGTGATGATACAGGCGATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132
 Qy 139 ATTCTCTGAGCTCTTTTGTTCATCGCGGCGACCTTGGGCGTGTGATCTTTTGTCCAAATGCC 198
 Db 133 ATCAGCGATTGGTCTGATTTGCGCGGATTTTGGTGGCAGGGGTTATGATGAG 192
 Qy 199 GCGCGCATGCTGCTGATTAATGCTGCTGGGCGGATCGCTTACCTGTTATGTTGCC 258
 Db 193 TCGCGGTGTTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 252
 Qy 259 GTCATGGCAGCGAAGAGCCCATGACAAACAGGTGGA 296
 Db 253 TTTGGCGCTTTTAAACAGCAATGAGCAATTAATGA 290

RESULT 11
 AAT33536/C
 ID AAT33536 standard; DNA; 15239 BP.
 XX
 AC AAT33536;
 XX
 DT 16-OCT-2003 (revised)
 DT 15-FEB-1998 (first entry)
 XX
 DE BCG deletion region 2 and flanking sequences.

XX BCG delta 2; virulence; avirulence; attenuation; gene deletion;
 KW mycobacteria; vaccine; infection; marker; ss.
 XX Mycobacterium bovis; strain BCG.

XX Key Location/Qualifiers
 FH 3382..14071
 FT misc_feature /tag= a
 FT /note= "BCG delta 1 deletion region"
 XX

EN WO9625519-A1.

XX 22-AUG-1996.

XX 15-FEB-1996; 96WO-US001938.

XX 17-FEB-1995; 95US-00390878.

XX (PATH-) PATHOGENESIS CORP.

XX Stover CK, Mahairas GG;

XX WPI; 1996-393419/39.

XX Detecting markers for avirulence in Mycobacterium - used in production of
 PT vaccines against bacterial infection, and to detect bacterial infection.

XX Example 1; Fig 2; 66pp; English.

XX This DNA sequence comprises Mycobacterium bovis BCG deletion sequence
 CC BCGdelta2. A specific genetic deletion of this region results in an
 CC avirulence phenotype of the mycobacterium. 2 Other deletion regions (see
 CC AAT33535 and AAT33537) have also been detected. Identification involved
 CC screening a BCG cosmid library with a radiolabeled probe obtained
 CC following DNA subtraction between virulent Mycobacterium tuberculosis
 CC H37Rv and avirulent BCG. The deletions provide useful markers for the
 CC identification of an avirulent, or a virulent, mycobacterial phenotype.
 CC Determination of avirulence requires the detection of the presence or
 CC absence of the deletion; the deletions are detected either by detecting
 CC the presence or absence of deletion junctions (see AAT33538-46), or by
 CC detecting the presence or absence of the sequences contained within the
 CC deletion. Deletion polypeptides are used as components of immunological
 CC assays and in vaccines. (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 15239 BP; 2878 A; 4791 C; 4729 G; 2841 T; 0 U; 0 Other;
 Query Match 8.0%; Score 57; DB 2; Length 15239;
 Best Local Similarity 53.1%; Pred. No. 4e-06;
 Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;
 Qy 431 AGCCCATGTTGATGCAATCGTCTGACCTGTTGAACCGCAATCGTATTGACGCGT 490
 Db 4792 AGTCTCTGTTGACTGTGCGGCATTCAGTTCTTCAACCAACATCTTACCTGACACCG 4733
 Qy 491 TTGTGTTTATCGCGGCGTTCGGCGCAATACCGCGACACCGGACGTTGATTTTCGCG 550
 Db 4732 TCGTGTGTTAGTGGCGCTGGCCACAGGACACGCGA---CCAGCGCTGCGTGTTCGCGC 4676
 Qy 551 CTGGCGCGTTCGGCGCAAGCTGATCTGTTCCGCTGCTGTTGGTTCGGCGACGACGAT 610
 Db 4675 TCGGCGCGCTCACAGCCAGTGGGTATGTTCCGCCACCTTCGGGTTTCGAGCGCGGT 4616
 Qy 611 TGTCAAGCCCGCTGTCCAGCCCAAGGTGTGGCGCTCGATCAACGCTCGTGGCAGTTG 670
 Db 4615 TCGCGCGGCTGTTCAACCAACCGCGCTGTGAGAAATCCTTCGACGCGCTGATCGCGTCA 4556
 Qy 671 TGATGACCGCATTTGCCCATCAAACTGATGTTGA 703
 Db 4555 TGATGTTGCGCTGGGAATCTCGTGCACCGTGA 4523

RESULT 12

AAT199682_22

Continuation (23 of 45) of AAT199682 from base 2200001 (Mycobacterium tuberculosis strai
 WP Sequence split into 45 fragments LOCUS AAT199682 Accession Aat199682

Fragment Name	Begin	End
WP AAT199682_00	1	110000
WP AAT199682_01	100001	210000
WP AAT199682_02	200001	310000
WP AAT199682_03	300001	410000
WP AAT199682_04	400001	510000
WP AAT199682_05	500001	610000
WP AAT199682_06	600001	710000
WP AAT199682_07	700001	810000
WP AAT199682_08	800001	910000
WP AAT199682_09	900001	1010000
WP AAT199682_10	1000001	1110000
WP AAT199682_11	1100001	1210000
WP AAT199682_12	1200001	1310000
WP AAT199682_13	1300001	1410000
WP AAT199682_14	1400001	1510000
WP AAT199682_15	1500001	1610000
WP AAT199682_16	1600001	1710000
WP AAT199682_17	1700001	1810000
WP AAT199682_18	1800001	1910000
WP AAT199682_19	1900001	2010000
WP AAT199682_20	2000001	2110000
WP AAT199682_21	2100001	2210000
WP AAT199682_22	2200001	2310000
WP AAT199682_23	2300001	2410000
WP AAT199682_24	2400001	2510000
WP AAT199682_25	2500001	2610000
WP AAT199682_26	2600001	2710000
WP AAT199682_27	2700001	2810000
WP AAT199682_28	2800001	2910000
WP AAT199682_29	2900001	3010000
WP AAT199682_30	3000001	3110000
WP AAT199682_31	3100001	3210000
WP AAT199682_32	3200001	3310000
WP AAT199682_33	3300001	3410000
WP AAT199682_34	3400001	3510000
WP AAT199682_35	3500001	3610000
WP AAT199682_36	3600001	3710000
WP AAT199682_37	3700001	3810000
WP AAT199682_38	3800001	3910000
WP AAT199682_39	3900001	4010000

30571 TGATGGTTGGGCTGGGAATCTCGCTGACCGTGA 30603

27870 TGATGGTTGCGCTGGGAATCTCGCTGACCGTGA 27902

CC poly

This invention relates to a novel method for identifying essential genes for survival or virulence of mycobacteria species. The method comprises aligning the genomic sequence of a first mycobacterium species on a genomic sequence of a second mycobacterium species and selecting a polynucleotide sequence that is highly conserved in both genomes with no

CC counterparts in other bacterial genomic sequences and that corresponds to
CC an essential gene for the survival or virulence of mycobacterium species.
CC The method of the invention is useful for detecting M. tuberculosis or M.
CC leprae infection. The method reduces the number of potential new targets
CC and protective antigens for new drugs and vaccine compositions to treat
CC and prevent mycobacterial diseases, particularly tuberculosis and
CC leprosy. The present sequence represents a Mycobacterial cosmid DNA
CC sequence used in the method of the invention
XX

SQ Sequence 86114 BP; 14693 A; 27103 C; 28984 G; 15334 T; 0 U; 0 Other;

Query Match 7.9%; Score 56.2; DB 6; Length 86114;
Best Local Similarity 53.2%; Pred. No. 1.6e-05;
Matches 143; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

QY 437 TCTTATGAGCAATCGTGTGACCTGCTGTTGAACCCGAAATCGTATTTGGACGCGTTTGTGT 496
DB 64559 TGCATAATGTGCTGTGTGTGACCTTTCTCAACCCACACAGTCTATCTGGACACTGTGTGT 64618
QY 497 TTATCGGGCGGCGCGCGCAATACGCGACACCGGAGGTGATTTTGGCGCGTGGCG 556
DB 64619 TGATCGGTGCGCTCGCCCAATAGGAATCAGA---TCTGGGTGTGTTCGAGCCGCGTG 64675
QY 557 CGTTCGCGCAAGCTGATCTGTTCCCGCTGCTGGTGGTTCGCGCAGCAGCATTTGTAC 616
DB 64676 CTTGGCGCGCGAGGCTGTATGTTTCGCGCTGTTGGATTTAGCGCGGCGGCTACAGC 64735
QY 617 GCCCGCTTCACGCCCAAGGTGTGGCGCTGGATCAACGCTGCTGGCGCGGCTGTGATGA 676
DB 64736 CATTTCTCGCAACTCCAGCTGTTGGCGCATCTCTGATGCGCTGTTGGCGTGACGATGA 64795
QY 677 CGCATTTGGCGCATCAACATGATTTGATG 705
DB 64796 TTGGGCTGCGCGTGTGTGTGCTGTACG 64824

RESULT 15
AAI99682_05
Continuation (6 of 45) of AAI99682 from base 500001 (Mycobacterium tuberculosis strain H
WP Sequence split into 45 fragments LOCUS AAI99682 Accession AAI99682

Fragment Name	Begin	End
WP AAI99682_00	1	110000
WP AAI99682_01	100001	210000
WP AAI99682_02	200001	310000
WP AAI99682_03	300001	410000
WP AAI99682_04	400001	510000
WP AAI99682_05	500001	610000
WP AAI99682_06	600001	710000
WP AAI99682_07	700001	810000
WP AAI99682_08	800001	910000
WP AAI99682_09	900001	1010000
WP AAI99682_10	1000001	1110000
WP AAI99682_11	1100001	1210000
WP AAI99682_12	1200001	1310000
WP AAI99682_13	1300001	1410000
WP AAI99682_14	1400001	1510000
WP AAI99682_15	1500001	1610000
WP AAI99682_16	1600001	1710000
WP AAI99682_17	1700001	1810000
WP AAI99682_18	1800001	1910000
WP AAI99682_19	1900001	2010000
WP AAI99682_20	2000001	2110000
WP AAI99682_21	2100001	2210000
WP AAI99682_22	2200001	2310000
WP AAI99682_23	2300001	2410000
WP AAI99682_24	2400001	2510000
WP AAI99682_25	2500001	2610000
WP AAI99682_26	2600001	2710000
WP AAI99682_27	2700001	2810000
WP AAI99682_28	2800001	2910000
WP AAI99682_29	2900001	3010000
WP AAI99682_30	3000001	3110000
WP AAI99682_31	3100001	3210000

WP AAI99682_32 3200001 3310000
WP AAI99682_33 3300001 3410000
WP AAI99682_34 3400001 3510000
WP AAI99682_35 3500001 3610000
WP AAI99682_36 3600001 3710000
WP AAI99682_37 3700001 3810000
WP AAI99682_38 3800001 3910000
WP AAI99682_39 3900001 4010000
WP AAI99682_40 4000001 4110000
WP AAI99682_41 4100001 4210000
WP AAI99682_42 4200001 4310000
WP AAI99682_43 4300001 4410000
WP AAI99682_44 4400001 4411529
Query Match 7.9%; Score 56.2; DB 4; Length 110000;
Best Local Similarity 53.2%; Pred. No. 1.8e-05;
Matches 143; Conservative 0; Mismatches 123; Indels 3; Gaps 1;
QY 437 TGTGATGGCAATCGTGTGACCTGCTGTTGAACCCGAAATCGTATTTGGACGCGTTTGTGT 496
DB 77960 TGCAATGTGCTGTGTGACCTTTCTCAACCCACACAGTCTATCTGGACACTGTGTGT 78019
QY 497 TTATCGGGCGGCGTGGCGCGCAATACGCGGACACCGGAGGTGATTTTGGCGCTGGCG 556
DB 78020 TGATCGGTGCGCTCGCCCAATAGGAATCAGA---TCTGGGTGTGTTCGAGCCGCGTG 78076
QY 557 CGTTCGCGCAAGCTGATCTGTTCCCGCTGCTGGTGGTTCGCGCAGCAGCATTTGTAC 616
DB 78077 CTTGGCGCGCAGGCTGTATGTTTCGCGCTGTTGGATTTAGCGCGGCGGCTACAGC 78136
QY 617 GCCCGCTGTCCAGCCCCAACAGGTGTGGCGCTGGATCAACGCTGCTGGCGAGTTTGTGATGA 676
DB 78137 CATTTCTCGCAACTCCAGCTGCTTGGCGCATCTCTGATGCGCTGTTGGCGTGACGATGA 78196
QY 677 CGCATTTGGCGCATCAACATGATTTGATG 705
DB 78197 TTGGGCTGCGCGTGTGTGTGCTGTACG 78225

Search completed: March 15, 2004, 14:59:14
Job time : 293.925 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:47:35 ; Search time 53.5608 Seconds
(without alignments)
7366.769 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1016_1726

Perfect score: 711

Sequence: 1 atgggtcatggaatctt.....aactgatgtgatgggttag 711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq*
- 5: /cgn2_6/ptodata/2/ina/PCUS_COMB.seq*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	61.8	8.7	699	4	US-09-252-991A-7908	Sequence 7908, Ap
2	61.8	8.7	834	4	US-09-252-991A-7643	Sequence 7643, Ap
3	61.8	8.7	894	4	US-09-252-991A-7841	Sequence 7841, Ap
4	57	8.0	597	4	US-09-894-844-11	Sequence 11, Appl
5	57	8.0	15239	1	US-08-390-878-17	Sequence 17, Appl
6	57	8.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
7	57	8.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
8	56.6	8.0	642	4	US-09-489-039A-4674	Sequence 4674, Ap
9	53.8	7.6	1176	4	US-09-252-991A-7571	Sequence 7571, Ap
10	43	6.0	645	4	US-09-543-681A-800	Sequence 800, App
11	39.4	5.5	505	4	US-09-621-976-15639	Sequence 15639, A
12	37.6	5.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
13	37.6	5.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl
14	35.4	5.0	636	4	US-09-328-352-1611	Sequence 1611, Ap
15	35	4.9	1002	4	US-09-252-991A-1099	Sequence 1099, Ap
16	35	4.9	1092	4	US-09-252-991A-946	Sequence 946, App
17	35	4.9	1521	4	US-09-252-991A-989	Sequence 989, App
18	34.8	4.9	723	4	US-09-252-991A-1130	Sequence 1130, Ap
19	34.8	4.9	1134	4	US-09-252-991A-1056	Sequence 1056, Ap
20	34.8	4.9	1233	4	US-09-252-991A-2121	Sequence 2121, Ap
21	34.8	4.9	1395	4	US-09-252-991A-904	Sequence 904, App
22	34.8	4.9	1029	4	US-09-252-991A-1954	Sequence 1954, App
23	34.6	4.9	1029	4	US-09-252-991A-1351	Sequence 1351, Ap
24	34.6	4.9	1029	4	US-09-252-991A-1752	Sequence 1752, Ap
25	34.6	4.9	1098	4	US-09-252-991A-1304	Sequence 1304, Ap
26	34.6	4.9	1215	4	US-09-252-991A-1975	Sequence 1975, Ap
27	34.6	4.9	1230	4	US-09-252-991A-2136	Sequence 2136, Ap

28	34.6	4.9	1632	4	US-09-252-991A-1216	Sequence 1216, Ap
29	34.6	4.9	1746	4	US-09-252-991A-1163	Sequence 1163, Ap
30	34.6	4.9	1995	4	US-09-252-991A-14545	Sequence 14545, A
31	34.6	4.9	3396	4	US-09-252-991A-14676	Sequence 14676, A
32	34.6	4.9	3444	4	US-09-252-991A-15078	Sequence 15078, A
33	34.6	4.9	6611	1	US-08-402-282-2	Sequence 2, Appl
34	34.6	4.9	6611	1	US-08-508-004-2	Sequence 2, Appl
35	34.6	4.9	6611	1	US-08-402-066-2	Sequence 2, Appl
36	34.6	4.9	6611	1	US-08-402-068-2	Sequence 2, Appl
37	34.2	4.8	495	4	US-09-252-991A-15856	Sequence 15856, A
38	34.2	4.8	804	4	US-09-252-991A-15809	Sequence 15809, A
39	34.2	4.8	963	4	US-09-489-039A-5986	Sequence 5986, Ap
40	34.2	4.8	2073	4	US-09-252-991A-15779	Sequence 15779, A
41	33.8	4.8	1284	4	US-09-252-991A-13377	Sequence 13377, A
42	33.8	4.8	1377	4	US-09-252-991A-13756	Sequence 13756, A
43	33.8	4.8	2318	4	US-09-620-312D-733	Sequence 733, App
44	33.6	4.7	717	4	US-09-252-991A-13565	Sequence 13565, A
45	33.6	4.7	1452	4	US-09-252-991A-10581	Sequence 10581, A

ALIGNMENTS

RESULT 1

US-09-252-991A-7908

; Sequence 7908, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 7908

; LENGTH: 699

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-7908

Query Match 8.7%; Score 61.8; DB 4; Length 699;

Best Local Similarity 54.6%; Pred. No. 4.9e-09;

Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;

Qy 440 TGATGGCAATCGTGTGCTGACCTGGTGTGACCCGAATGCGTATTGGACGGTTGTGTTA 499

Db 434 TGGCGGCACTGGCGGTGACCTGCTCAACCCCACTCTATCTCGATACCGPACTTCTGA 493

Qy 500 TCGCGCGGCTGGCGGCAATACCGGACACCGGACGATTTTCGCCGCTGGCGCGT 559

Db 494 TCGGTTCTGGCGGCGCCACGAGG-----CGCGCGGCGGCTATGCCCTCGCGCGG 547

Qy 560 TCGCGGCAACCTGATCTGTGTTCCGCTGGTGGTTTCGGCGAGCAGCATTTGACGCC 619

Db 548 CCAGCGCTCGCTGATGTTCTTTCGCCCTCGCCCTCGCGCGGCAATGGCTGGCGCCCT 607

Qy 620 CGCTGTTCAGCCCAAGGTGTGGCGCTGGATCAAGTGTCTGGGAGTTGTGATGACCG 679

Db 608 GCGTGGCGCGCGCGCCACCTGGCGGCTCTGCTGACGTATGTTGGCGGCGCATGATGCTGG 667

Qy 680 CATTGGCCATCAACTGATGTTGATGGTTA 710

Db 668 GCATGGCGGCGCACTGCTGTTCCGGGATA 698

RESULT 2

US-09-252-991A-7643/c

; Sequence 7643, Application US/09252991A

Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7643
LENGTH: 834
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7643

Query Match 8.7%; Score 61.8; DB 4; Length 834;
Best Local Similarity 54.6%; Pred. No. 5.3e-09;
Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;
QY 440 TGATGGCAATCGTGTGACCTGGTTGAACCCGAATGCGTATTTCGACGGCTTGTGTTA 499
DB 386 TGGCGGCACTGGCGGTACCCCTGCTCAACCCACGCTCTATCTCGATACCGTACTTCTGA 327
QY 500 TCGCGGGCGTGGCGGCGCAATACGGCGACACCGGACGGTGGATTTCGCGCGTGGCGGT 559
DB 326 TCGGTTGCTGCGCGCCCGCAGG-----CCGCGCGGGCGCTATGCCCCCTCGGCGCG 273
QY 560 TCGCGGCAACCTGATCTGCTGCGGTGCGGTGCGGCGGCGAGCATGTTCAGCGCC 619
DB 272 COAGGCTCGTGTGATGTTCTTCGCCCTGCGCGCGGCGGATGGCTGGCCCCCT 213
QY 620 CGCTGTCCAGCCCGCAAGTGTGGCGTGGATCAAGCTGCTGGCAGTGTGTGATGACCG 679
DB 212 GCGTGGCGCGCGCGCCACCTGGCGCTGCTGCGACCTGATGTGGCGGCGCATGTGCTG 153
QY 680 CATTTGGCCATCAACTGATGTTGATGGTTA 710
DB 152 GCATGGCGCGCAACTGCTGTTCGGGGGATA 122

RESULT 3
US-09-252-991A-7841
; Sequence 7841, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7841
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7841

Query Match 8.7%; Score 61.8; DB 4; Length 894;
Best Local Similarity 54.6%; Pred. No. 5.5e-09;
Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;
QY 440 TGATGGCAATCGTGTGACCTGGTTGAACCCGAATGCGTATTTCGACGGCTTGTGTTA 499
DB 339 TGGCGGCACTGGCGGTACCCCTGCTCAACCCCGCAGTCTAATCTGATACCGTACTTCTGA 398

QY 500 TCGCGGCGGTTCGCGCGCAATACGGCGACACCGGACGGTGGATTTCGCGCGCTGGCGCGT 559
DB 399 TCGGTTGCTGCGCGCCCGCAGG-----CCGCGCGGGCGCTATGCGCTTCGCGCGCG 452
QY 560 TCGCGGCAACCTGATCTGCTGCGGTTCGCGGTTCGCGCGGCGGCGAGCATGTTCAGCGC 619
DB 453 CCAGCGCGCTGCTGATGTTGTTCTTCGCCCTCGCGCGGCGCATGGCTGGCCCCCT 512
QY 620 GCGTGTCCAGCCCGCAAGTGTGGCGCTGATCAAGCTGCTGCGGCGAGTGTGTGATGACCG 679
DB 513 GCGTGGCGCGCGCGGCGACCTGGCGCTGCTGACCTGATGTGTGGCGGCGCATGTGCTGG 572
QY 680 CATTTGGCCATCAACTGATGTTGATGGTTA 710
DB 573 GCATGGCGCGCAACTGCTGTTCGGGGGATA 603

RESULT 4
US-09-894-844-11
; Sequence 11, Application US/09894844
; Patent No. 6686166
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-09-894-844-11

Query Match 8.0%; Score 57; DB 4; Length 597;
Best Local Similarity 53.1%; Pred. No. 1.3e-07;
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;
QY 431 AGCCCATGTTGATGGCAATCGTGTGCTGACCTGGTTGAACCGAATGCGTATTTCGACGGT 490
DB 326 AGGTCTCTGTGACCTGTCGGGCAATTCAGTTCTCAACCCACAGCTCTACTCGACACG 385
QY 491 TTGTGTTTATCGCGCGGTTCGCGCGCAATACCGCGACACCGGACCGTGGATTTCGCGCG 550
DB 386 TCGTGTTCGTAGCGCGCTGGCCAAACGACGACAGCA---CCAGCGCTGGCTGTTCGCGCC 442
QY 551 CTGCGCGGTTCGCGCGGACGCTGATCTGGTTCCGCTGGTGGTTTCGCGCGGCGAGCAT 610
DB 443 TCGCGCGGTTCACAGCGGAGTGGCGGTATGTTCGCCACCTCGGGTTCGGAGCGCGCGGT 502
QY 611 TGTACAGCGCGCTGCTCCAGCGCCCAAGGTGTGGCGCTGGATCAAGCTGCTGTCGCGAGTTG 670
DB 503 TCGCGCGGTGTTTCAACACCGCGCTGCTGGAGATCTCTGACGGCTGATCGCGGTCA 562
QY 671 TGATGACCGCATTTGGCCATCAAACTGATGTTGA 703
DB 563 TGATGTTGCGTGGGAATCTGCTGACCGGTGA 595

RESULT 5
US-08-390-878-17/c
; Sequence 17, Application US/08390878
; Patent No. 5700683

GENERAL INFORMATION:

APPLICANT: Scover, Charles K.
APPLICANT: Mahairas, Gregory G.
TITLE OF INVENTION: VIRULENCE-ATTENUATING GENETIC DELETIONS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourille and Crew
STREET: One Market Plaza, Steuart Street Tower, 20th
STREET: Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/390,878

FILING DATE: 17-FEB-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hunter, Tom

REGISTRATION NUMBER: 38,498

REFERENCE/DOCKET NUMBER: 15371A-17

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/543/9600

TELEFAX: 415/543/5043

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 15239 base pairs

TYPE: nucleic acid

STRADEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-390-878-17

Query Match

Best Local Similarity 8.0%; Score 57; DB 1; Length 15239;

Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 431 AGCCCATGTTGATGGCAATCGTCTGACCTGTTGAACCCGGAATCGGTATTGGACGGT 490
DB 4792 AGTCTCTGTGACTGTGCGGCAATTCAGTTCTCAACCCACACGCTTACCTCGACCG 4733
QY 491 TTGTGTTATCGCGCGGTGCGCGCGCAATACGCGGACACCGGACGTTGCGCG 550
DB 4732 TCGTGTGTTAGCGCGCTGGCCAAACGAGACAGCGA---CCAGCGCTGCTGTTCCGCC 4676
QY 551 CTGGCGGTTCCGCGGAGCTGATCTGTTCCCGCTGTTGCGGCGACGAGCAT 610
DB 4675 TCGGCGGTTCCAGCCAGTGGCGGTATGTTTCGCCACCTCGGGTTGGAGCGCGCGT 4616
QY 611 TGTACGCGCTGTTCCAGCGCCCAAGGTGTGGCGCTGATCAACGTCGTGGCGAGTTG 670
DB 4615 TGGCGGGGCTGTTCAACCAACCGCGCTGTGGAGAACTCTCGACGCGCTGATCGCGTCA 4556
QY 671 TGATGACCGCATTTGCCATCAAACTGATGTTGA 703
DB 4555 TGATGTTGCGTGGGAATCTCGCTGACCGTGA 4523

RESULT 6

US-09-103-840A-2

Sequence 2, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 2

LENGTH: 4403765

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

FEATURE:

OTHER INFORMATION: CDC 1551

OTHER INFORMATION: "n" bases at various positions throughout the sequence

OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match

Best Local Similarity 8.0%; Score 57; DB 3; Length 4403765;

Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 431 AGCCCATGTTGATGGCAATCGTCTGACCTGTTGAACCCGGAATCGGTATTGGACGGT 490
DB 2227633 AGTCTCTGTGACTGTGCGGCAATTCAGTTCTCAACCCACACGCTTACCTCGACCG 2227692
QY 491 TTGTGTTATCGCGCGGTGCGCGCGCAATACGCGGACACCGGACGTTGCGCG 550
DB 2227693 TCGTGTGTTAGCGCGCTGGCCAAACGAGACAGCGA---CCAGCGCTGCTGTTCCGCC 2227749
QY 551 CTGGCGGTTCCGCGGAGCTGATCTGTTCCCGCTGTTGCGGCGACGAGCAT 610
DB 2227750 TCGGCGGTTCCAGCCAGTGGCGGTATGTTTCGCCACCTCGGGTTGGAGCGCGT 2227809
QY 611 TGTACGCGCTGTTCCAGCGCCCAAGGTGTGGCGCTGATCAACGTCGTGGCGAGTTG 670
DB 2227810 TGGCGGGGCTGTTCAACCAACCGCGCTGTGGAGAACTCTCGACGCGCTGATCGCGTCA 2227869
QY 671 TGATGACCGCATTTGCCATCAAACTGATGTTGA 703
DB 2227870 TGATGTTGCGTGGGAATCTCGCTGACCGTGA 2227902

RESULT 7

US-09-103-840A-1

Sequence 1, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 1

LENGTH: 4411529

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match

Best Local Similarity 8.0%; Score 57; DB 3; Length 4411529;

Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 431 AGCCCATGTTGATGGCAATCGTCTGACCTGTTGAACCCGGAATCGGTATTGGACGGT 490
DB 2230334 AGTCTCTGTGACTGTGCGGCAATTCAGTTCTCAACCCACACGCTTACCTCGACCG 2230393
QY 491 TTGTGTTATCGCGCGGTGCGCGCGCAATACGCGGACACCGGACGTTGCGCG 550

```

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7571
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7571

Query Match
; Sequence 4674, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4674
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4674

```

RESULT 8

```

; Sequence 4674, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4674
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4674

```

```

Query Match
; Sequence 4674, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4674
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4674

```

RESULT 9

```

; Sequence 7571, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

```

```

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7571
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7571

```

```

Query Match
; Sequence 4674, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4674
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4674

```

RESULT 10

```

; Sequence 800, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 800
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-800

```

```

Query Match
; Sequence 800, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 800
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-800

```

```

RESULT 12
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:

```

Query Match	5.3%	Score 37.6;	DB 3;	Length 4411529;
Best Local Similarity	47.5%;	Pred. No. 8.7;		
Matches 112;	Conservative 0;	Mismatches 124;	Indels 0;	Gaps 0;
QY	430	AAGCCCATGTTTGATGGCAATCTGTGCTGACCTGTGGTTGAACCCGAATGCGTATTTGGACGGC	489	
DB	719010	AAGGACCTTGCCTTGCGCGTTCAGCGGGCGCGGCGGAGCTGCCCGGGCCCCCGTGGT	718951	
QY	490	TTTGTGTTTATCGCGGGCGCTCGGCGCGCAATACGGCAGACCCGGAACGGTGTGGATTTTTCGCC	549	
DB	718950	GGCGTGTCTTTCACCTGCAACCGGCGGGGACGACGAATGTTTCGGGGTTCACCGACCCAGC	718981	
QY	550	GCTGGGGGTTTCGGCGCAAGCTGATCTGTGTTCCCGCTGTGTGGTTCGGCGCGCAGCAGCA	609	
DB	718990	GCGTCAGATGAGGACCTGCTGGCGGGATTCCTGCTGGCCGGTTCTTCGCCCGCGGG	718931	
QY	610	TTGTACGCCCGCTGTGTCAGCCCAAGGTGTGGCGCTGCGATCAACCTGCTGTGCGC	665	
DB	718930	GAGATCGGCCCGGTTCGGGGGCCACAAACGCTTGACACGGGTTTACCGCGTCGATGGC	718775	
RESULT 14				
US-09-328-1611				
Sequence 1611, Application US/09328352				

	Query Match	4.9%	Score 35;	DB 4;	Length 1002;
	Best Local Similarity	48.3%	Pred. No. 0.83;		
	Matches 131;	Conservative	0;	Mismatches 135;	Indels 5; Gaps 1;
QY	420	GGTTTGCGTAAAGCCATGTTGATGCGCAATCGTGTGACCTGGTGTGAACCCGGAATCGTA	479		
Db	613	GGCTTTGATCGAGCGGAGCGAGAGGTCTATCTGCCAGGTCCATCGAGTCGAAGAGAA	672		
QY	480	TTTGAGCGCGTTTGTGTTTATCGCGCGCGTTCGGCGCGCATACGGACACCGGACGGTG	539		
Db	673	CGCCAGCGCGATGATCAGAAAGACAGTCGCTGATACGACTCAGCGGACGCGTTCCAG	732		
QY	540	GAATTTCCGCGTCGCGCGTTCGCGCAAGCCTGATCTGGTTCCGCTGTGGTGGTTCGCG	599		
Db	733	GGGTTTCGCGCGGATACGCGCGCGTFCGGGCACTGACTGGTTCCTCTGCGGGCGCGGCGT	792		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2004, 00:51:03 ; Search time 254.891 Seconds
(without alignments)
10269.549 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1016_1726

Perfect score: 711
Sequence: 1 atgttgatcaggaatctt.....aactgatgttgaggtag 711

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 184079884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 - 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
 - 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	711	100.0	711	14	US-10-166-142-7
2	711	100.0	822	10	US-09-746-660A-51
3	711	100.0	2374	14	US-10-196-232-24
4	711	100.0	3309400	9	US-09-738-626-1
5	708	99.6	708	9	US-09-738-626-3455
6	696.8	98.0	712	14	US-10-166-142-9
7	62.8	8.8	609	14	US-10-156-761-1109
8	62.8	8.8	9025608	14	US-10-156-761-1
9	57	8.0	597	9	US-09-894-844-11
10	57	8.0	597	15	US-10-388-902-11
11	56.2	7.9	86114	14	US-10-080-170-648
12	42.2	5.9	3000	14	US-10-156-761-5949
13	42.2	5.9	9025608	14	US-10-156-761-1
14	40.4	5.7	1392	15	US-10-369-493-31368
15	40.4	5.7	1443	15	US-10-369-493-28610

c 16	39.4	5.5	594	15	US-10-369-493-33277	Sequence 33277, A
c 17	39.2	5.5	1257	14	US-10-156-761-5999	Sequence 5999, Ap
c 18	39	5.5	993	10	US-09-746-660A-47	Sequence 47, Appl
c 19	39	5.5	2823	12	US-10-282-122A-26243	Sequence 26243, A
c 20	39	5.5	2826	12	US-10-282-122A-28428	Sequence 28428, A
c 21	38.6	5.4	1245	12	US-10-282-122A-14346	Sequence 14346, A
c 22	38	5.3	1018	12	US-10-425-114-960	Sequence 960, App
c 23	38	5.3	1458	14	US-10-156-761-1306	Sequence 1306, Ap
c 24	38	5.3	2457	12	US-10-425-114-28221	Sequence 28221, A
c 25	38	5.3	2457	12	US-10-425-114-30614	Sequence 30614, A
c 26	37.8	5.3	536	16	US-10-338-110-119	Sequence 119, App
c 27	37.8	5.3	2038	14	US-10-094-113-11	Sequence 11, Appl
c 28	37.6	5.3	536	16	US-10-338-110-119	Sequence 119, App
c 29	37.6	5.3	2079	14	US-10-156-761-1755	Sequence 1755, Ap
c 30	37.6	5.3	3073	12	US-10-424-599-33204	Sequence 33204, A
c 31	37.6	5.3	3090	12	US-10-425-114-11304	Sequence 11304, A
c 32	37.4	5.3	624	14	US-10-156-761-2439	Sequence 2439, Ap
c 33	37.4	5.3	1218	15	US-10-369-493-44172	Sequence 44172, A
c 34	37.4	5.3	1350	12	US-10-282-122A-15180	Sequence 15180, A
c 35	37.4	5.3	1631	15	US-10-369-493-33258	Sequence 33258, A
c 36	37.2	5.2	424	9	US-09-960-352-12392	Sequence 12392, A
c 37	37.2	5.2	873	12	US-10-282-122A-25710	Sequence 25710, A
c 38	37	5.2	945	12	US-10-282-122A-14462	Sequence 14462, A
c 39	37	5.2	1410	14	US-10-156-761-3506	Sequence 3506, Ap
c 40	37	5.2	1464	12	US-10-282-122A-31748	Sequence 31748, A
c 41	36.8	5.2	546	14	US-10-156-761-4089	Sequence 4089, Ap
c 42	36.8	5.2	678	12	US-10-282-122A-26365	Sequence 26365, A
c 43	36.8	5.2	681	12	US-10-282-122A-28183	Sequence 28183, A
c 44	36.8	5.2	954	12	US-10-282-122A-26045	Sequence 26045, A
c 45	36.6	5.1	681	15	US-10-260-238-3605	Sequence 3605, Ap

ALIGNMENTS

RESULT 1
US-10-166-142-7
; Sequence 7, Application US/10166142
; Publication No. US20030124687A1
; GENERAL INFORMATION:
; APPLICANT: GUNJI, YOSHIYA
; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
; FILE REFERENCE: 223789US
; CURRENT APPLICATION NUMBER: US/10/166,142
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001-1777075
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(711)
; OTHER INFORMATION:
US-10-166-142-7

Query Match	100.0%	Score 711;	DB 14;	Length 711;
Best Local Similarity	100.0%	Pred. No. 3.7e-225;		
Matches 711;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGTTGATCATCGAAATCTTCATTACAGTCTGCTTTTGGGGCCAGCTCTTTTACTGTCC	60	
Db	1	ATGTTGATCATCGAAATCTTCATTACAGTCTGCTTTTGGGGCCAGCTCTTTTACTGTCC	60	
Qy	61	ATCGACCCGACAGATGTACTGTGATTAACAAGAAATTAAGCGGAGGACTCATTTGCG	120	
Db	61	ATCGACCCGACAGATGTACTGTGATTAACAAGAAATTAAGCGGAGGACTCATTTGCG	120	

121 GTTCTTCTGCTGTTAAATTTCTGACGCTCTTTTGTTCATCGCCGACACCTGGGCGTT 180
121 GTTCTTCTGCTGTTAAATTTCTGACGCTCTTTTGTTCATCGCCGACACCTGGGCGTT 180
181 GATCTTTTGTCCAAATGCGCGCGATGCTGCTGCTGATATATGCGCTGGGGTGGCATCGCT 240
181 GATCTTTTGTCCAAATGCGCGCGATGCTGCTGCTGATATATGCGCTGGGGTGGCATCGCT 240
241 TACTGTTATGTTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
241 TACTGTTATGTTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
301 CCACAGATCAATGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGGCGTTGCT 360
301 CCACAGATCAATGAAGAAACAGAACCAACCGTGGCCGATGACACGCTTTGGGCGTTGCT 360
361 GCGGTGGCCATCTGACACGCGCAACCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 420
361 GCGGTGGCCATCTGACACGCGCAACCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 420
421 GTTTGGGTAAGCCCATGTTGATGGCAATGCTGCTGACCTGTTGAACCCGAATGCGTAT 480
421 GTTTGGGTAAGCCCATGTTGATGGCAATGCTGCTGACCTGTTGAACCCGAATGCGTAT 480
481 TTGGACCGCTTTGTTTATCGCGCGCTGCGCGCGCAATAACCGGACACCGGACGCTG 540
481 TTGGACCGCTTTGTTTATCGCGCGCTGCGCGCGCAATAACCGGACACCGGACGCTG 540
541 ATTTTCCGCGCTGGCGGCTGCGCGCGCAACCGGCTGGGCTGGGCTGGGCTGGGCTGGG 600
541 ATTTTCCGCGCTGGCGGCTGCGCGCGCAACCGGCTGGGCTGGGCTGGGCTGGGCTGGG 600
601 GCACGACATTTGTCACGCGCGCTGTCACGCGCGCAACCGGCTGGGCTGGGCTGGGCTGG 660
601 GCACGACATTTGTCACGCGCGCTGTCACGCGCGCAACCGGCTGGGCTGGGCTGGGCTGG 660
661 GTGGCAGTTGTGATACCGGCTTTGGCCATCAAACTGATGTTGATGGGTTAG 711
661 GTGGCAGTTGTGATACCGGCTTTGGCCATCAAACTGATGTTGATGGGTTAG 711

RESULT 2

US-09-746-660A-51
; Sequence 51, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08

; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 51
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(799)
; OTHER INFORMATION: RXA01394
US-09-746-660A-51

Query Match 100.0%; Score 711; DB 10; Length 822;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTGATCATGCAATCTCTTACAGTCTGCTTTTGGGCGCCAGTCTTTTACTGCTCC 60
Db 92 ATGCTGATCATGCAATCTCTTACAGTCTGCTTTTGGGCGCCAGTCTTTTACTGCTCC 151
Qy 61 ATCGGACCGCAGAAATGATCTGCTGATTTAAACAAAGGAATTAAGCGCGAAGACTCATTTGG 120
Db 152 ATCGGACCGCAGAAATGATCTGCTGATTTAAACAAAGGAATTAAGCGCGAAGACTCATTTGG 211
Qy 121 GTTCTTCTGCTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCCGACACCTGGGCGTT 180
Db 212 GTTCTTCTGCTGTTTAAATTTCTGACGCTCTTTTGTTCATCGCCGACACCTGGGCGTT 271
Qy 181 GATCTTTTGTCCAAATGCGCGCGCTGCTGCTGATATATGCGCTGGGCTGGCATCGCT 240
Db 272 GATCTTTTGTCCAAATGCGCGCGCTGCTGCTGATATATGCGCTGGGCTGGCATCGCT 331
Qy 241 TACTGTTATGTTTGGCTCATCGCAGCGAAAGACGCGATGACAAACAAAGTGGAGCG 300
Db 332 TACTGTTATGTTTGGCTCATCGCAGCGAAAGACGCGATGACAAACAAAGTGGAGCG 391
Qy 301 CCACAGATCAATGAAGAAACAGAACCAACCGTGGCGCAATGACGCTTTGGGCGTTGCT 360
Db 392 CCACAGATCAATGAAGAAACAGAACCAACCGTGGCGCAATGACGCTTTGGGCGTTGCT 451
Qy 361 GCGGTGGCCATCTGACACGCGCAACCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 420
Db 452 GCGGTGGCCATCTGACACGCGCAACCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 511
Qy 421 GTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 512 GTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571
Qy 481 TTGGACCGCTTTGTTTATCGCGCGCTGCGCGCGCAATGACGCGACACCGGACGCTGG 540
Db 572 TTGGACCGCTTTGTTTATCGCGCGCTGCGCGCGCAATGACGCGACACCGGACGCTGG 631
Qy 541 ATTTTCCGCGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 632 ATTTTCCGCGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691
Qy 601 GCGGACATTTGTCACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 692 GCGGACATTTGTCACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 751
Qy 661 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 711
Db 752 GTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTTAG 802

RESULT 3

US-10-196-232-24
; Sequence 24, Application US/10196232
; Publication No. US20030113899A1
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, MIKIKO
; APPLICANT: ITO, HISAO
; APPLICANT: GUNJI, YOSHIYA

```

; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: 225391US0
; CURRENT APPLICATION NUMBER: US/10/196,232
; CURRENT FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: JP 2001-224586
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 2374
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1025)..(1723)
; OTHER INFORMATION:
US-10-196-232-24

Query Match      100.0%; Score 711; DB 14; Length 2374;
Best Local Similarity 100.0%; Pred. No. 6,7e-225;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGATCATGAAATCTTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTC 60
DB 1016 ATGTGATCATGAAATCTTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTC 1075
QY 61 ATCGACCCGAGATGTAAGTAAACAAAGAAATTAAGCGCGAAGGACTCATTTGG 120
DB 1076 ATCGACCCGAGATGTAAGTAAACAAAGAAATTAAGCGCGAAGGACTCATTTGG 1135
QY 121 GTTCTTCTGCTGTTTAAATTTCTGAGTCTTTTGTTCATCGCGGACCTTGGGGCTT 180
DB 1136 GTTCTTCTGCTGTTTAAATTTCTGAGTCTTTTGTTCATCGCGGACCTTGGGGCTT 1195
QY 181 GATCTTTTGTCCAAATGCGCGGCTGCTGCATATATATGCTGGGGTGGCATCGCT 240
DB 1196 GATCTTTTGTCCAAATGCGCGGCTGCTGCATATATATGCTGGGGTGGCATCGCT 1255
QY 241 TACTGTATGTTTGGCTCATGCGAGGAAAGACCCATGACAAACAAAGTGGAGCG 300
DB 1256 TACTGTATGTTTGGCTCATGCGAGGAAAGACCCATGACAAACAAAGTGGAGCG 1315
QY 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGCCCGATGACACGCTTTGGGGCTT 360
DB 1316 CCACAGATCATTTGAAGAAACAGAACCAACCGTGCCCGATGACACGCTTTGGGGCTT 1375
QY 361 GCGTGGCCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGCATAAGCAGCG 420
DB 1376 GCGTGGCCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGCATAAGCAGCG 1435
QY 421 GTTTGGGTAAAGCCATTTGATGCGAATCGTGTGACCTGGTTGAACCCGATGCGTAT 480
DB 1436 GTTTGGGTAAAGCCATTTGATGCGAATCGTGTGACCTGGTTGAACCCGATGCGTAT 1495
QY 481 TTGACGCGTTTGTGTTTATCGCGCGCTGCGCGCAATACCGGACACCGGAGCGTG 540
DB 1496 TTGACGCGTTTGTGTTTATCGCGCGCTGCGCGCAATACCGGACACCGGAGCGTG 1555
QY 541 ATTTTCCGCGCTGGCGCTTCCGCGCAAGCCTGATCTGTTCCCGCTGGTGGTTCCGC 600
DB 1556 ATTTTCCGCGCTGGCGCTTCCGCGCAAGCCTGATCTGTTCCCGCTGGTGGTTCCGC 1615
QY 601 CGACGACATTTGACGCGCTGTCAGCCCAAGGTGTGGCGTGTGATCAACGTCGTC 660
DB 1616 CGACGACATTTGTCAGCGCTGTCAGCCCAAGGTGTGGCGTGTGATCAACGTCGTC 1675
QY 661 GTGCGAGTTGATGACCGCATTTGGCGCATCAACTGATGTTGATGGTTAG 711
DB 1676 GTGCGAGTTGATGACCGCATTTGGCGCATCAACTGATGTTGATGGTTAG 1726

```

RESULT 4

```

US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match      100.0%; Score 711; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 2.4e-223;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGATCATGAAATCTTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTC 60
DB 1328953 ATGTGATCATGAAATCTTACAGTCTGCTTTGGGGCCAGTCTTTTACTGTC 1328894
QY 61 ATCGGACCCGAGATGTAAGTAAACAAAGAAATTAAGCGCGAAGGACTCATTTGG 120
DB 1328893 ATCGGACCCGAGATGTAAGTAAACAAAGAAATTAAGCGCGAAGGACTCATTTGG 1328834
QY 121 GTTCTTCTGCTGTTTAAATTTCTGAGTCTTTTGTTCATGCGCGGACCTTGGGGCTT 180
DB 1328833 GTTCTTCTGCTGTTTAAATTTCTGAGTCTTTTGTTCATGCGCGGACCTTGGGGCTT 1328774
QY 181 GATCTTTTGTCCAAATGCGCGGCTGCTGCATATATGCTGGGGTGGCATCGCT 240
DB 1328773 GATCTTTTGTCCAAATGCGCGGCTGCTGCATATATGCTGGGGTGGCATCGCT 1328714
QY 241 TACTGTATGTTTGGCTCATGCGAGGAAAGACCCATGACAAACAAAGTGGAGCG 300
DB 1328713 TACTGTATGTTTGGCTCATGCGAGGAAAGACCCATGACAAACAAAGTGGAGCG 1328654
QY 301 CCACAGATCATTTGAAGAAACAGAACCAACCGTGCCCGATGACACGCTTTGGGGCTT 360
DB 1328653 CCACAGATCATTTGAAGAAACAGAACCAACCGTGCCCGATGACACGCTTTGGGGCTT 1328594
QY 361 GCGTGGCCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGCATAAGCAGCG 420
DB 1328593 GCGTGGCCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTGCATAAGCAGCG 1328534
QY 421 GTTTGGGTAAAGCCATTTGATGCGAATCGTGTGACCTGGTTGAACCCGATGCGTAT 480
DB 1328533 GTTTGGGTAAAGCCATTTGATGCGAATCGTGTGACCTGGTTGAACCCGATGCGTAT 1328474
QY 481 TTGACGCGTTTGTGTTTATCGCGCGCTGCGCGCAATACCGGACACCGGAGCGTG 540
DB 1328473 TTGACGCGTTTGTGTTTATCGCGCGCTGCGCGCAATACCGGAGCGTG 1328414
QY 541 ATTTTCCGCGCTGGCGCTTCCGCGCAAGCCTGATCTGTTCCCGCTGGTGGTTCCGC 600

```

Db 1328413 ATTTTCGCGCTGGCGGTTCGGCGCAAGCTGATCTGTTTCGCGTGGTTCGGC 1328354
 QY 601 GCACGACATTTGACGCGCGCTCTCCAGCCCAAGGTGGCGCTGGATCAAGTGTCT 660
 Db 1328353 GCACGACATTTGACGCGCGCTCTCCAGCCCAAGGTGGCGCTGGATCAAGTGTCT 1328294
 QY 661 GTGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 711
 Db 1328293 GTGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAG 1328243

RESULT 5
 US-09-738-626-3455
 ; Sequence 3455, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENO, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OKAZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 3455
 ; LENGTH: 708
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3455

Query Match 99.6%; Score 708; DB 9; Length 708;
 Best Local Similarity 100.0%; Pred. No. 3.7e-224;
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGTGATCATGGAATCTTTACAGGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60
 Db 1 ATGGTGATCATGGAATCTTTACAGGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60
 QY 61 ATCGACCGCAGATGCTGCTGATTAACAGGAATTAAGCCGAGGACTCATTTGCG 120
 Db 61 ATCGACCGCAGATGCTGCTGATTAACAGGAATTAAGCCGAGGACTCATTTGCG 120
 QY 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGTCTTTTTCATCGCCGACCTTTGGCGGT 180
 Db 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGTCTTTTTCATCGCCGACCTTTGGCGGT 180
 QY 181 GATCTTTTGTCCATGCGCGCGATCGTGTGATTAACAGGAATTAAGCCGAGGACTCATTTGCG 240
 Db 181 GATCTTTTGTCCATGCGCGCGATCGTGTGATTAACAGGAATTAAGCCGAGGACTCATTTGCG 240
 QY 241 TACTCTTATGTTTGGCGTCAATGCGGAGGAGCCATGACAAAGAGGTGGAAGCG 300
 Db 241 TACTCTTATGTTTGGCGTCAATGCGGAGGAGCCATGACAAAGAGGTGGAAGCG 300
 QY 301 CCACAGATCATTAAGAAACAGAACCAACCGTGCCTGATGACAGCGCTTTGGCGGTTTCG 360
 Db 301 CCACAGATCATTAAGAAACAGAACCAACCGTGCCTGATGACAGCGCTTTGGCGGTTTCG 360

QY 361 GCGTGGCCACTGACACGCGCAACCGGTCGCGGTGGAGTGAGCTCGATAAGCAGCG 420
 Db 361 GCGTGGCCACTGACACGCGCAACCGGTCGCGGTGGAGTGAGCTCGATAAGCAGCG 420
 QY 421 GTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGTTGACCCGATCGGTAT 480
 Db 421 GTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGTTGACCCGATCGGTAT 480
 QY 481 TTGGAACGCGTGTGTTTATCGCGCGCTCGCGCGCAATACGGCCGACACCGGACGTTGG 540
 Db 481 TTGGAACGCGTGTGTTTATCGCGCGCTCGCGCGCAATACGGCCGACACCGGACGTTGG 540
 QY 541 ATTTTCGCGCTGGCGCTTCGCGCAAGCTGATCTGCTCCGCTGGTGGTTCGCGC 600
 Db 541 ATTTTCGCGCTGGCGCTTCGCGCAAGCTGATCTGCTCCGCTGGTGGTTCGCGC 600
 QY 601 GCACGACATTTGACGCGCGCTTCAGCCGCCCAAGGTGGCGCTGCGATCAACGTCTC 660
 Db 601 GCACGACATTTGACGCGCGCTTCAGCCGCCCAAGGTGGCGCTGCGATCAACGTCTC 660
 QY 661 GTGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGT 708
 Db 661 GTGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGT 708

RESULT 6
 US-10-166-142-9
 ; Sequence 9, Application US/10166142
 ; Publication No. US20030124687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GUNJI, YOSHIYA
 ; APPLICANT: YASUEDA, HISASHI
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
 ; FILE REFERENCE: 223789US
 ; CURRENT APPLICATION NUMBER: US/10/166,142
 ; CURRENT FILING DATE: 2002-06-11
 ; PRIOR APPLICATION NUMBER: JP 2001-1777075
 ; PRIOR FILING DATE: 2001-06-12
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 9
 ; LENGTH: 712
 ; TYPE: DNA
 ; ORGANISM: Brevibacterium lactofermentum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(375)
 ; OTHER INFORMATION:
 US-10-166-142-9

Query Match 98.0%; Score 696.8; DB 14; Length 712;
 Best Local Similarity 99.6%; Pred. No. 1.9e-220;
 Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 1 ATGGTGATCATGGAATCTTTACAGGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60
 Db 1 ATGGTGATCATGGAATCTTTACAGGTCTGCTTTTGGGGCCAGTCTTTTACTGTCC 60
 QY 61 ATCGACCGCAGATGCTGCTGATTAACAGGAATTAAGCCGAGGACTCATTTGCG 120
 Db 61 ATCGACCGCAGATGCTGCTGATTAACAGGAATTAAGCCGAGGACTCATTTGCG 120
 QY 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGTCTTTTTCATCGCGGACCTTTGGCGGT 180
 Db 121 GTTCTTCTCGTGTGTTTAAATTTCTGACGTCTTTTTCATCGCGGACCTTTGGCGGT 180
 QY 181 GATCTTTTGTCCATGCGCGCGATCGTGTGATTAACAGGAATTAAGCCGAGGACTCATTTGCG 240
 Db 181 GATCTTTTGTCCATGCGCGCGATCGTGTGATTAACAGGAATTAAGCCGAGGACTCATTTGCG 240
 QY 241 TACTCTTATGTTTGGCGTCAATGCGGAGGAGCCATGACAAAGAGGTGGAAGCG 300

```
Db 241 TACCTGTATTGTTTCCGTCATGCGCAGCAAGAACGCCATGACAAAACAGGTGGAAGCG 300
QY 301 CCACAGATCATTTGAAGAAACAGAACCAACCTGCGCCGATGACACGCTTTGGGCG-GTTC 359
Db 301 CCACAGATCATTTGAAGAAACAGAACCAACCTGCGCCGATGACACGCTTTGGGCGGTTC 360
QY 360 GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTCGATGAAGCAGCG 419
Db 361 GCGGTGGCCACTGACACGCGCAACCGGGTGGAGGTGAGCGTCGATGAAGCAGCG 420
QY 420 GGTTCGGTAAAGCCATGTTGATGCAATCGTCTGACCTGTTGAACCCGATGCGTA 479
Db 421 GGTTCGGTAAAGCCATGTTGATGCAATCGTCTGACCTGTTGAACCCGATGCGTA 480
QY 480 TTTGGACGGTTTGTGTTATCGCGCGGTGCGCGCGCAATACGGGACACCGGAGCGTG 539
Db 481 TTTGGACGGTTTGTGTTATCGCGCGGTGCGCGCGCAATACGGGACACCGGAGCGTG 540
QY 540 GATTTTCGCGCTGCGCGCTTCGCGCAAGCTGATCTGTTCCGCGTGGTGGTTTCGG 599
Db 541 GATTTTCGCGCTGCGCGCTTCGCGCAAGCTGATCTGTTCCGCGTGGTGGTTTCGG 600
QY 600 CGCAGCAGATTTGACGCGCGCTGTCAGCGCCCAAGGTGTCGCGCTGATCAACCTGCT 659
Db 601 CGCAGCAGATTTGACGCGCGCTGTCAGCGCCCAAGGTGTCGCGCTGATCAACCTGCT 660
QY 660 CTTGGCAGTTGTGATGACCGCATTTGCCATCAAACTGATGTTGATGGTTAG 711
Db 661 CTTGGCAGTTGTGATGACCGCATTTGCCATCAAACTGATGTTGATGGTTAG 712

RESULT 7
US-10-156-761-1109
; Sequence 1109, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1109
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(609)
US-10-156-761-1109

Query Match 8.8%; Score 62.8; DB 14; Length 609;
Best Local Similarity 52.2%; Pred. No. 4.7e-10;
Matches 164; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

QY 385 CGGGTCGGGTGGAGTGACGCTGATTAAGCAGCGGGTTGGGTAAAGCCCATGTTGATG 444
Db 289 CCGGCGGGCTGCGGGCGGAGGCGCGCGGAGGGCTCCCGCAGACGCGCGCTACTCACC 348
QY 445 GCAATCGTCTGACCTGTTGAACCGCAATGCTATTGACGCGCTTGTGTTATCGGC 504
Db 349 TGCTGGCCCTTACCTGGCTCAACCCGACGCTTACTCTGACAC---CGTCTCTCTCTC 405

Query Match 8.8%; Score 62.8; DB 14; Length 609;
Best Local Similarity 52.2%; Pred. No. 4.7e-10;
Matches 164; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

QY 385 CGGGTCGGGTGGAGTGACGCTGATTAAGCAGCGGGTTGGGTAAAGCCCATGTTGATG 444
Db 289 CCGGCGGGCTGCGGGCGGAGGCGCGCGGAGGGCTCCCGCAGACGCGCGCTACTCACC 348
QY 445 GCAATCGTCTGACCTGTTGAACCGCAATGCTATTGACGCGCTTGTGTTATCGGC 504
Db 349 TGCTGGCCCTTACCTGGCTCAACCCGACGCTTACTCTGACAC---CGTCTCTCTCTC 405
```

```
QY 505 GCGCTCGGCGCAATACGGCAGCAGCGGATTTTCGCCCTCGGCGGCTTCGCG 564
Db 406 GCGCTCGATCGCGCGCGACCGCGGCTCACTGCGCTGACCTTCGGCTCTCGCGCGGCGCTC 465
QY 565 GCAAGCCTGATCTGCTCCCGCTGCGGTGCGCGCGCAGCAGCATTTGTCACGCGCGCTG 624
Db 466 GCGAGCCTGCTGCTGCTGCGCGCGCTCGCGCTTCGCGCTCGCGCGCTTCCTG 525
QY 625 TCCAGCCCCAAGGTGTCGCGCTGATCAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 684
Db 526 GCGCGCGCGCTCGCGCATGCGCGGTGCTGAGCGCGCTGCTGCGCGCGCATGTCACGATG 585
QY 685 GCCATCAAACTGAT 698
Db 586 GCGCGCACCTGAT 599

RESULT 8
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (418715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 8.8%; Score 62.8; DB 14; Length 9025608;
Best Local Similarity 52.2%; Pred. No. 5.5e-08;
Matches 164; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

QY 385 CGGGTCGGGTGGAGTGACGCTGATTAAGCAGCGGGTTGGGTAAAGCCCATGTTGATG 444
Db 1406101 CCGCGCGGCTGCGGGCGGAGGCGCGCGGAGGGCTCCCGCAGACGCGCGCTACTCACC 1406042
QY 445 GCAATCGTCTGACCTGTTGAACCGCAATGCTATTGACGCGCTTGTGTTATCGGC 504
Db 1406041 TGCTGGCCCTTACCTGGCTCAACCCGACGCTTACTCTGACACC---GTCTTCTCTCTC 1405985
QY 505 GCGCTCGGCGCAATACGGCAGCAGCGGTGATTTTCGCCCTGGGCGGCTTCGCG 564
Db 1405984 GCGTTCGATCGGCGCGCAGCGCGGCTCACTGCGCTGACCTTCGGTCTCGCGCGGCGCTC 1405925
QY 565 GCAAGCCTGATCTGCTCCCGCTGCGGTGCGGTTCGCGCGCAGCAGCATTTGTCACGCGCGCTG 624
Db 1405924 GCGAGCCTGCTGCTGCTGCGCGCGCTTCGCGCTTCGCGCTCAACGCTGCTCGCGCGCTTCCTG 1405865
QY 625 TCCAGCCCCAAGGTGTCGCGCTGATCAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 684
Db 1405864 GCGCGCGCGCTGCGCATGCGCGGTGCTGAGACGCCCTGCTGCGCGCGCATGTCACGATG 1405805
QY 685 GCCATCAAACTGAT 698
```

Db 1405804 GCGGCCACCCCTGAT 1405791

RESULT 9

US-09-894-844-11
; Sequence 11, Application US/09894844
; Patent No. US20020176873A1
; GENERAL INFORMATION.

```

/ APPLICANT: Benji, Michael
/ APPLICANT: Small, Peter
/ APPLICANT: Schoolnik, Gary
/ APPLICANT: Wilson, Michael A.
/ TITLE OF INVENTION: Molecular Differences Between Species of
/ TITLE OF INVENTION: the M. Tuberculosis Complex
/ FILE REFERENCE: STAN102CON
/ CURRENT APPLICATION NUMBER: US/09/894,844
/ CURRENT FILING DATE: 2001-06-27
/ PRIOR APPLICATION NUMBER: 09/318,191
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: 60/097,936
/ PRIOR FILING DATE: 1998-08-25
/ NUMBER OF SEQ ID NOS: 137
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11
/ LENGTH: 597
/ TYPE: DNA
/ ORGANISM: Mycobacteria tuberculosis
US-09-894-844-11

```

Query Match	8.0%;	Score 57;	DB 9;	Length 597;
Best Local Similarity	53.1%;	Pred. No. 3.9e-08;		
Matches 145;	Conservative 0;	Mismatches 125;	Indels 3;	Gaps 1;
QY	431	AGCCCATTTGATGGCAATCGTGTGACTGCTTGTGAACCCGAATGCTATTGGACGCGT	490	
Db	326	AGGTCCTGGTGACCTGTGGGCAATTACGTTCTCTCAACCCACACACGCTTACCTGCACACGG	385	
QY	491	TTGTGTTTATCGGCGCGCTCGGCGCGGCAATACGGCGACACCGGACGGTGGATTTCGCGG	550	
Db	386	TGCTGTGTCTAGCGCGCTTGGCCAAAGAGACAGCGA---CCAGCGCTGGCTGTTCGCGC	442	
QY	551	CTGGGCGCTTTCGGGCGACGCTGATCTGTGTTCCCTCGCTGGTGGGTTTCGGGCGACAGCAT	610	
Db	443	TCGGCGGCTACAGCCAGTGGCGTATGTTTCGCCACCTTCGGGTTCGAGCGCGCCGCT	502	
QY	611	TGTCAGCGCGCTGTTCAGCGCCGAAAGTGTGGCGCTGGATCAACGTCGTCTGGACGATTG	670	
Db	503	TGGCGGGCTGTTTCCACAAACCCCGCTCTGGTGGAGAACTCTCGACGCGCTGATCGCGGTCA	562	
QY	671	TGATGACCGCAATTGGCCATCAAACTGATTTGA	703	
Db	563	TGATGTTTGCCTGGGAATCTTCGTGACCGTGA	595	

RES. III, T. 10

US-10-388-902-11 ; Sequence 11, Application US/10388902
; Publication No. US20040018574A1
; GENERAL INFORMATION.

```

; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-10-388-902-11

Query Match      8.0%; Score 57; DB 15; Length 597;
Best Local Similarity 53.1%; Pred. No. 3.9e-08;
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

Qy  431  AGCCCATGTTGATGGCAATGCTCTGACCTGGTTGAACCGGAATGCGTATTTGGACCGGT 490
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  326  AGGCTCGGTGTGACCTGTGTGGGCATTTACAGTTCTTCAACACACAGTCTACCTCGACACG 385
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  491  TTGCTGTTTATCGGCGGCGCTCGGGCGGCAATACGGCGACACACCGACCGTGGATTTTCGCCG 550
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  386  TCGTGTGCTAGGCGGCTGSCCAACGACACACGCA---CCAGCGCTGGCTGTTTCGCGC 442
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  551  CTGGCGGCTTCGGGGCAAGCCGTGATCTGGTTCCGCGTGGTGGGTTTCGGGGCAGCAGCAT 610
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  443  TCGGCGGCTGCACAGCCAGTGGCGTATGGTTTCGCCACCCCTCGGGTTCCGAGCGCGCGGT 502
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  611  TGTACAGCCCGCTGTTCAGGCCCAAGGTGTGGCTGGGATCAAGTGTGCTGGCGCAGTTG 670
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  503  TCGCGGGCTGTTTCACCAACCCGGCTGTGGAGATTCCTCGACGGCTGTATCGCGGTCA 562
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  671  TGATGACCGCATGGGCATCAAACTGATTTGA 703
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  563  TGATGTTGGCTGGGAATCTCGCTGACCGTGA 595
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-10-080-170-648
; Sequence 648, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S. T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR

```

RESIT.T 11

```

RESULT 11
US-10-080-170-648
; Sequence 648, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOME
; IDENTIFYING TARGETS FOR THE
; TREATMENT OF MYCOBACTERIOSES
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495, 0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 648
; LENGTH: 86114
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-648

```

Query Match	7.9%;	Score 56.2;	DB 14;	Length 86114;
Best Local Similarity	53.2%;	Pred. No. 8.5e-07;		
Matches 143;	Conservative 0;	Mismatches 123;	Indels 3;	Gaps 1;
Qy	437	TGTTGATGGCAATCGGTGCTGACTGTTTGAAACCGAATCGGTATTTCGACGCGTTTGTTGTT	496	
Db	64559	TGCAAAATGTCCTGGTGGTGACCTTTCTCAACCCACACGCTCTATCTGGACACTGTGGTGT	64618	
Qy	497	TTATCGGCGCGCTCGCGCGGCATACGGCGACACCGGACGGTGGATTTTCGCGCGCTGCGC	556	
Db	64619	TGATCGGTGGCCCTCGCCCAATAGGAATACGA---TTCGGGTGGTTTTTCGGAGCCGGTG	64675	
Qy	557	CGTTCGGGCAAGCCTGATCTGTGTTCCGCTGGTGGTTTTCGGCGCAGACGACATGTTAC	616	
Db	64676	CCTGGCGCGCCGCGTTCGTATGGTTTTCGCGTGTGGGATTTAGCGGGCGCGCTACAGC	64735	

QY 617 GCGCGTGTCCAGCCCAAGGTGGCGTGGATCAAGTCGTGGCGAGTTGTGATGA 676
 Db 64736 CATTTTCGCACTCACTGCTTGGCGCATCTTTGATGCGCTGGTTCGCGTACGATGA 64795
 QY 677 CCGCATGGCCATCAAACTGATGTGATG 705
 Db 64796 TTGGGGTCCGCGCTGTGTGTCTGTCTCAG 64824

RESULT 12

US-10-156-761-5949
 ; Sequence 5949, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 5949
 ; LENGTH: 3000
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(3000)
 US-10-156-761-5949

Query Match 5.9%; Score 42.2; DB 14; Length 3000;
 Best Local Similarity 51.9%; Pred. No. 0.0072;
 Matches 123; Conservative 0; Mismatches 108; Indels 6; Gaps 1;
 QY 369 CACTGACACGCGCAACCGGGTGGCGGTGGAGTGGATCGATAGCAGCGGGTTGGGT 428
 Db 57 CACCGATCCGTGGCGCGCGAGCGGTGCTGGAGAGCGCGAGCTGGAGCCCGTCAAGGC 116
 QY 429 AAAGCCCATGTTGATGGCAATCTGCTGACCTGTTGAACCCGCAATGCGTATTTGGACGC 488
 Db 117 CGACCCCGTGTCTGTGGAAGCCCTG-----GGCGGACCGCGCATCCCGATCTGGCGCT 170
 QY 489 GTTTGTGTTTATCGCGCGCTGCGCGCAATACGCGACACCGGACGGTGGATTTTCGC 548
 Db 171 GCTCGACTGCTCCCGCTGTGCGAGCGGAGAGACACACCGCCCGCGGAGTCTCT 230
 QY 549 CGCTGGCGGGTTCGCGGCAAGCTGATCTGTTCCCGCTGGTGGGTTCGCGCGCAGC 605
 Db 231 CGACACGCTGATCGCGGCCAAACCCCTGCGGACCGGCTCTCTCGGCGTCTCTCGGTGC 287

RESULT 13

US-10-156-761-1
 ; Sequence 1, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 1
 ; LENGTH: 9025608
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (4187715)
 ; OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1
 Query Match 5.9%; Score 42.2; DB 14; Length 9025608;
 Best Local Similarity 51.9%; Pred. No. 0.38;
 Matches 123; Conservative 0; Mismatches 108; Indels 6; Gaps 1;
 QY 369 CACTGACACGCGCAACCGGGTGGCGGTGGAGTGGATCGATAGCAGCGGGTTGGGT 428
 Db 7187526 CACCGATCCGTGGCGCGCGAGCGGTGCTGGAGAGCGCGAGCTGGAGCCCGTCAAGGC 7187585
 QY 429 AAAGCCCATGTTGATGGCAATCTGCTGACCTGTTGAACCCGCAATGCGTATTTGGACGC 488
 Db 7187586 CGACCCCGTGTCTGTGGAAGCCCTG-----GGCGGACCGCGCATCCCGATCTGGCGCT 7187639
 QY 489 GTTTGTGTTTATCGCGCGCTGCGCGCAATACGCGACACCGGACGGTGGATTTTCGC 548
 Db 7187640 GCTCGACTGCTCCCGCTGTGCGAGCGGAGAGACACACCGCCCGCGGAGCTCTCT 7187699
 QY 549 CGCTGGCGGGTTCGCGGCAAGCTGATCTGTTCCCGCTGGTGGGTTCGCGCGCAGC 605
 Db 7187700 CGACACGCTGATCGCGGCCAAACCCCTGCGGACCGGCTCTCTCGGCGTCTCTCGGTGC 7187756

RESULT 14

US-10-369-493-31368
 ; Sequence 31368, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 31368
 ; LENGTH: 1392
 ; TYPE: DNA
 ; ORGANISM: Burkholderia cepacia
 US-10-369-493-31368
 Query Match 5.7%; Score 40.4; DB 15; Length 1392;
 Best Local Similarity 51.1%; Pred. No. 0.019;
 Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
 QY 417 GCGGGTTTGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGTGGTGAACCGGAATGC 476
 Db 793 GCGGTGGCGGTGCGCGGGAAGTTTCAGACCTCGGGGCAAGATTGTTCGCGGCGAATCG 842
 QY 477 GTATTGGACGGCTTTGTGTTTATCGCGCGGCTCGGCGGCAATACGGCGACACCGGAGC 536

Db 843 CATTTCCTCGCGCGCGCTGTACGAAGCGTTCGTCGCCCAATACAGCGAGCGCGTCAA 902
QY 537 GTGGATTTCCTCGCGCGCTGTACGAAGCGTTCGTCGCCCAATACAGCGAGCGCGTCAA 902
Db 903 GCGGCTGGAAGTTCGCGCGCTGTACGAAGCGTTCGTCGCCCAATACAGCGAGCGCGTCAA 902
QY 597 CGCGCG 602
Db 963 AGCCGC 968

RESULT 15

US-10-369-493-28610
; Sequence 28610, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIORITY FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIORITY FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 28610
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Burkholderia fungorum
US-10-369-493-28610

Query Match 5.7%; Score 40.4; DB 15; Length 1443;
Best Local Similarity 51.1%; Pred. No. 0.02; Mismatches 91; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 417 GCGGTTTGGTAAAGCCCATGTTGATGGCAATCTGTGTGACCTGTGTAACCCGGAATGC 476
Db 825 GCGGTTGCGGCTCGCGCGAGTTTCAGACCTCGGGGCGAGGATTGTTGCGCGCGGAATCG 884
QY 477 GTATTTCGACGCGTTTGTGTTATCGCGCGGCTCGCGCGGCAATACGGGACACCGGAG 536
Db 885 CATTTCGTCGCGCGCGCTGTACGAAGCGTTCGTCGCCCAATACAGCGAGCGCGTCAA 944
QY 537 GTGGATTTCCTCGCGCGCTGTACGAAGCGTTCGTCGCCCAATACAGCGAGCGCGTCAA 944
Db 945 GCGGCTGGAAGTTCGCGCGCTGTACGAAGCGTTCGTCGCCCAATACAGCGAGCGCGTCAA 944
QY 597 CGCGCG 602
Db 1005 AGCCGC 1010

Search completed: March 17, 2004, 09:46:58
Job time : 282.891 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:43:00 ; Search time 1955.25 Seconds
(without alignments)
10858.975 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1016_1726

Perfect score: 711

Sequence: 1 atgtgtatcatggaatctt.....aactgatgttgatgggttag 711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_estum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_esti:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estom:*

16: em_gss_hum:*

17: em_gss_inv:*

18: em_gss_pln:*

19: em_gss_vrt:*

20: em_gss_fun:*

21: em_gss_mam:*

22: em_gss_mus:*

23: em_gss_pro:*

24: em_gss_rtd:*

25: em_gss_phg:*

26: em_gss_vrl:*

27: gb_gss1:*

28: gb_gss2:*

29: gb_gss3:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	60.4	8.5	830	29	CNS01MQH
C 2	56.8	8.0	1034	28	BZ554294
C 3	54.2	7.6	897	28	BZ560550
C 4	51.4	7.2	1620	28	BZ568946

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C	5	48.6	6.8	833	28	BZ556963
	6	48.6	6.8	925	29	CNS0091P
	7	47.6	6.7	852	13	BX393687
	8	46	6.5	614	14	CA830247
	9	46	6.5	720	14	CD881613
	10	44.4	6.2	620	14	CA829022
	11	44.4	6.2	644	13	BQ704105
	12	42.8	6.0	532	10	AM927407
	13	42.8	6.0	538	14	CD861975
	14	41.8	5.9	885	13	BX425603
	15	41.2	5.8	910	29	CNS0060N
	16	40.8	5.7	605	13	CA127255
	17	40.8	5.7	614	13	CA141725
	18	40.8	5.7	619	14	CA192339
	19	40.8	5.7	619	14	CA238619
	20	40.2	5.7	601	14	CF134261
	21	40.2	5.7	613	10	BE498441
	22	40	5.6	486	13	BZ668006
	23	40	5.6	656	14	CA595136
	24	40	5.6	754	14	CB619745
	25	40	5.6	781	28	BZ989800
	26	40	5.6	798	14	CB624901
	27	40	5.6	839	29	CG446626
	28	40	5.6	925	29	CNS0091P
	29	39.8	5.6	569	13	BZ62199
	30	39.8	5.6	672	13	BZ324760
	31	39.8	5.6	676	13	BZ296506
	32	39.8	5.6	729	13	BZ033325
	33	39.8	5.6	737	13	BZ034198
	34	39.8	5.6	1201	13	BX332962
	35	39.6	5.6	1667	11	AV108197
	36	39.2	5.5	647	14	CA284378
	37	39	5.5	566	14	CA189386
	38	39	5.5	680	13	CA114181
	39	39	5.5	950	11	AY104298
	40	38.8	5.5	557	10	AW787854
	41	38.8	5.5	609	14	CF015635
	42	38.8	5.5	725	14	CA185740
	43	38.6	5.4	624	14	CA263962
	44	38.4	5.4	1201	13	BX361080
	45	38.2	5.4	711	13	BZ79094

ALIGNMENTS

RESULT 1
CNS01MQH/c
LOCUS
DEFINITION
CNS01MQH 830 bp DNA linear GSS 14-JUN-2001
Anopheles gambiae GSS T7 end of clone 22E24 of NotreDamel library
from strain PEST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AL151258
AL151258.1 GI:7011737
GSS.
Anopheles gambiae (African malaria mosquito)
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
Submitted (16-FEB-2000) BEWI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the

1 (bases 1 to 830)
Genoscope.
Direct Submission
Submitted (16-FEB-2000) BEWI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the

2 (bases 1 to 830)
Genoscope.
Direct Submission
Submitted (16-FEB-2000) BEWI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the

1 (bases 1 to 830)
Genoscope.
Direct Submission
Submitted (16-FEB-2000) BEWI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
Submitted (16-FEB-2000) BEWI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the

Result No.	Score	Query Match	Length	DB ID	Description
C 1	60.4	8.5	830	29	CNS01MQH
C 2	56.8	8.0	1034	28	BZ554294
C 3	54.2	7.6	897	28	BZ560550
C 4	51.4	7.2	1620	28	BZ568946

Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.

FEATURES

source
1. .830
Location/Qualifiers
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PBST"
/db_xref="taxon:7165"
/clone="22824"
/clone_lib="NotreDamel"
/note="end : T7"

ORIGIN

Query Match 8.5%; Score 60.4; DB 29; Length 830;
Best Local Similarity 50.7%; Pred. No. 9.5e-05;
Matches 142; Conservative 1; Mismatches 137; Indels 0; Gaps 0;

QY 23 TTACAGCTCTGCTTTGGGGCCAGCTTTTACTGTCATCGGACCGCAGAAATGACTCG 82
Db ATCAAGGGCTTGCTTACGTCATTTTATTTTCCCTCGCGCAAAATCGTTG 538
QY 83 TGATTAACAAGGAATTAAGCGCAAGGACTCATTTGGGTTCTTCTGTTGTTTATTT 142
Db TGATGATCAGGCAATCGCGCCAGTACCATCTGATGATGCTTATGCGGGTAA 478
QY 143 CTGAGCTCTTTTGTTCATCGCGCACTTGGCGTTGATCTTTTGTCCAAATCGCGCG 202
Db GTGATTTGCTGCTGATCTGTCTCGGATTTTGGCGGCGGCTGCTGATGAGTCTC 418
QY 203 CGATCGTCTCGCATATTAATGCTGGGGTGGCATCGCTTACCTGTTATGTTTCCGTCA 262
Db CTGCTGCTGCTGGGTTGTGTCACCTGGGGCGGCTGGGCTTCTGCTGTTACGATTGG 358
QY 263 TGGCAGCGAAGACGCCATGACAAACAAGTGGAGGCGCC 302
Db GCGGCTGAAGCGCGATGAGCAGTAACCTCGAACTGGC 318

RESULT 2

BZ554294
LOCUS
DEFINITION
pacsl-60.4617.x1 pacsl-60 Pseudomonas aeruginosa genomic clone
pacsl-60.4617, genomic survey sequence.

ACCESSION
BZ554294
VERSION
BZ554294.1 GI:27161466
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.

REFERENCE

1 (bases 1 to 1034)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1. .1034
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacsl-60.4617"
/clone_lib="pacsl-60"
/note="Clinical isolate 1-60 Whole genomic shotgun"

ORIGIN

Query Match 7.6%; Score 54.2; DB 28; Length 897;
Best Local Similarity 52.8%; Pred. No. 0.0041;
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;
QY 440 TGATGGCAATCTGCTGACCTGGTTGACCCCAATGCTATTGGACGCGTTTGTATTA 499
Db 552 TCGGGGCACTGGCGTTACCTGGCTTAACCCCAACGCTATTTCGATACCGTACTTCTAC 493

ORIGIN

Query Match 8.0%; Score 56.8; DB 28; Length 1034;
Best Local Similarity 50.4%; Pred. No. 0.0009;
Matches 139; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 7 ATCATGGAATCTTCATTACAGTCTGCTTTTGGGGCCAGTCTTTTACTGTCCATCGGA 66
Db AGTTTCAGAGCTATCTCAACGGCATCTTGTTGGTGGCGGCTGTATCATCGCCATCGGC 141
QY 67 CCGCAGAATGCTACTGCTGATTAAACAAGGAATTAAGCGCGAAGGACTCATTTGGGTTCTT 126
Db GCGCAGAATGCTTCTGCTCGCAGAGCTGCGCGCGAGCATCACTCTCGGTAGCC 201
QY 127 CTCGTGTTTAAATTTCTGAGCTCTTTTGTTCATCGCGGACCTTTGGGCGTTGATCTT 186
Db GCGCTCTGCTGTTCTGCGAGCGGCTGCTGTCAAGCTCGGCGGTTTCGCGCTGSCCAAG 261
QY 187 TTGTCAATCGCGCGCATGCTGCTGATATTATGCGTGGGGTGGCATCGCTTACTG 246
Db CTGCTGCTGAAACCCGAGCTGCTGCGCATCGCCCTCGGGCGGAATCGCCTTCTCTG 321
QY 247 TTATGTTTCCGCTCATGCGCAGGAGAGCGCATG 282
Db ACCTGGTACGGGCTCAAGGCGCTCTTCGCGGCTTG 357

RESULT 3

BZ560550/c
LOCUS
DEFINITION
pacsl-164.2569.x1 pacsl-164 Pseudomonas aeruginosa genomic clone
pacsl-164.2569, genomic survey sequence.

ACCESSION
BZ560550
VERSION
BZ560550.1 GI:27178749
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.

REFERENCE

1 (bases 1 to 897)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1. .897
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacsl-164.2569"
/clone_lib="pacsl-164"
/note="Clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

Query Match 7.6%; Score 54.2; DB 28; Length 897;
Best Local Similarity 52.8%; Pred. No. 0.0041;
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;
QY 440 TGATGGCAATCTGCTGACCTGGTTGACCCCAATGCTATTGGACGCGTTTGTATTA 499
Db 552 TCGGGGCACTGGCGTTACCTGGCTTAACCCCAACGCTATTTCGATACCGTACTTCTAC 493

```

QY 500 TCGCGCGCTCGCGCGCAATACGCGCACCGGATGTTTCGCGCTGCGCGT 559
Db 492 GGTTCGCTCGCGCGCAGCAAGCGCG-----CGGGCGNTATGCCCTCGCGCGG 441
QY 560 TCGCGGCAACCTGATCTGTTCCGCTGCTGTTGCGCGCAGCAGATTCACGCC 619
Db 440 CAAAGCGCTCGCTGATGTTCTTCGCGCTCGCGCGCGCATGCTGCGCGCT 381
QY 620 CGCTGTCACAGCCCAAGGTGCGGCTGGATCAAGCTGCTGCGCAGTTGATGACCG 679
Db 380 GCTGCGCGCGCGCGCGCTGCGCGCTGCTGACCTGATGTTGGCGCGCATGATGCTG 321
QY 680 CATTGGCCATCAAACTGATGTTGATGGTTA 710
Db 320 GCATGGCGCGCAACTGCTGTTCCGGGATA 290

RESULT 4
BZ568946/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 1620)
AUTHORS
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL
J. Bacteriol. (2002) In press
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..1620
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."

FEATURES
source
Query Match 7.2%; Score 51.4; DB 28; Length 1620;
Best Local Similarity 54.3%; Pred. No. 0.027;
Matches 100; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 527 ACACCGGAGGTGGATTTTCGCGCTGCGCGGCTGCGCGCAGCTGATGTTCCCGC 586
Db 618 AGNCCGCGGCGCGCATATGCCCTCGCGCGCGCAGCGCTCGATGTTGTTCTCG 559
QY 587 TGGTGGGTTTCGCGCGCAGCAGCATTTGACGCGCGCTGTCCAGCCCAAGGTGTGCGCT 646
Db 558 CCCTGCGCTTCGCGCGCGCATGCTGCGCGCTGCTGCGCGCGCGCGCATGCTGCGCGC 499
QY 647 GATCAACCTGCTGCGCGCATGTTGATGACCGCATTCGCGCATCAAACTGATGTTGATGG 706
Db 498 TGCTGACATGATGTTGCGCGCGCATGCTGCTGCGCGCATGCTGCTGCTGCTGCTGCTG 439
QY 707 GTTA 710

```

```

Db 438 GATA 435

RESULT 5
BZ556963/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 833)
AUTHORS
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL
J. Bacteriol. (2002) In press
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..833
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone_lib="pacs1-60"
/note="clinical isolate 1-60 Whole genomic shotgun
library."

ORIGIN
Query Match 6.8%; Score 48.6; DB 28; Length 833;
Best Local Similarity 50.6%; Pred. No. 0.12;
Matches 117; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 52 TTATGTCATCGGACCGCGCAGATGCTGCTGATTAACAGAGATTAAGCGCGAAGGA 111
Db 446 TCATCGGCGCATCGCGCGCGCAGATGCTTCTCGCGCGAGAGCCTCGCGCGAGCAT 387
QY 112 CTCATTGCGGTTCTTCTCGTGTGTTAAATTCGACGCTCTTTTGTTCATCGCGCGCACC 171
Db 386 CACCTCTCGGTAGCGCGCTCTGCGTGTCTCGGACAGGCTGCTGAGCTCGGCGTG 327
QY 172 TGGGCGTTGATGATTTTGTCTCAATGCGCGCGCATGCTGCTCGATATTATGCGCTGGGGT 231
Db 326 TTCGCGCTGCGCGCAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 267
QY 232 GGCATGCTTACCTGTTATGTTTTCGCTGATGCGGCGCGCGCGCGCGCGCGCGCG 282
Db 266 GGGATCGCTCTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216

RESULT 6
CNS0091P
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC19D16 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)

```

ORGANISM

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 925)

AUTHORS

Direct Submission

TITLE

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)

JOURNAL

- Web : www.genoscope.cns.fr

COMMENT

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosawa and
Aaron Mamoser in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
21 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 6.8%; Score 48.6; DB 29; Length 925;
Best Local Similarity 13.8%; Pred. No. 0.12;
Matches 49; Conservative 157; Mismatches 149; Indels 0; Gaps 0;
QY 325 CCAACGTCGCGATGACAGCGCTTGGGGGTTGGCGGTGGCCATGACAGCGCAAC 384
DB 571 SC 630
QY 385 CGCGTCGCGGTGAGCGTTCGATAGCAGCGCGTTTGGTAAAGCCCATGTTGATG 444
DB 631 CSSSSSCSSSSSTSSSTSSSTSSSSSSSSSSSSSTSSSTSSSSSSSSSSSSSS 690
QY 445 GCATCGTGTGACCTGTTGAAACCGAATGCGTATTGACGCGTTTGTGTTATCGGC 504
DB 691 STGTSSTSSSSSSSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 750
QY 505 GCGTCGCGCGCAATACGCGACACGCGAGCGTATTCGCGCGTGGCGGTTCCGC 564
DB 751 SSSSSSTSCCTCCSSSSSSSSSTSSSSSTSSSSSSSSSSSSSSSSSSSSSSSS 810
QY 565 GCAAGCCTGATCTGTTTCCCGTGGTGGGTTTCGGCGCAGCAGCATTTGACCGCCG 624
DB 811 MCTCCSTYMBCTSYTSCGSSSSSSSSSGGVTGCGCGCGSSSTNGMTSSSSSSSC 870
QY 625 TCACCCCAAGGTGTGCGCGCTGATCAACGCTGCTGCGCAGTTGTGATGACCG 679
DB 871 SSS 925

RESULT 7

EX393687
LOCUS
DEFINITION BX393687 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC001YG07 5-PRIME, mRNA sequence.
ACCESSION BX393687
VERSION BX393687.1 GI:30624032
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 852)

AUTHORS

Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DC001AD04QP1.

FEATURES

Location/Qualifiers
1..852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC001YG07"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 6.7%; Score 47.6; DB 13; Length 852;
Best Local Similarity 17.2%; Pred. No. 0.22;
Matches 69; Conservative 152; Mismatches 178; Indels 3; Gaps 1;
QY 23 TTACAGTCTGCTTTGGGGGCGAGTCTTTTACTGTCATCGACCGCAGCAATGACTGG 82
DB 381 TTTTGGKTK 440
QY 83 TGATTAAACAGGAATTAAGCGCGAAGCACTCAATTCGGTCTTCTGCTGTGTTAATT 142
DB 441 KKK 500
QY 143 CTGACGCTTTTGTTCATCGCGGCGACCTTGGCGGCTTGTATCTTTGTCATGCGCGC 202
DB 501 KDKKK 560
QY 203 CGATCGTCTGCTGATATTATTCGCTGGGCTGGCATCGCTTACCTGTTATGTTGCGCTCA 262
DB 561 CTTTCKKK 620
QY 263 ---TGGCAGCGAAGAGCGCGCATGACAAACAGGTGGAAGCGCCACAGATCATTAAGAA 319
DB 621 KTTTCKKK 680
QY 320 CAGAACCAACGCTGCCCGATGACAGCCCTTTTGGCGGCTTGGCGGCTGCGCCACTGACAGC 379
DB 681 CGGMMVVVVVMMGMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 740
QY 380 GCAACCGGCTGCGGCTGAGCGTTCGATTAAGCAGAGCGGG 421
DB 741 MWMWMWMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMM 782

RESULT 8

CA830247/c
LOCUS
DEFINITION CA830247 614 bp mRNA linear EST 12-DEC-2002
1117004E09.y1 1117 - Unigene V from Maize Genome Project Zea mays
cDNA, mRNA sequence.
ACCESSION CA830247
VERSION CA830247.1 GI:26558012
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1 (bases 1 to 614)
AUTHORS
TITLE
JOURNAL
COMMENT

Maize ESTs from various cDNA libraries sequenced at Stanford

Walbot, V.
University
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1117004 row: E column: 09.

FEATURES
source

Location/Qualifiers
1..614
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="dbEST:946110C04.y1"
/db_xref="taxon:4577"
/clone_lib="1117 - Unigene V from Maize Genome Project"

/note="This library represents the unique genes found in the fifth round of EST sequencing at Stanford University for the maize genome project. Sequences are present from library 946. Contigs were assembled using ZmDBAssembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."

ORIGIN

Query Match 6.5%; Score 46; DB 14; Length 614;
Best Local Similarity 46.3%; Pred. No. 0.51;
Matches 151; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 349 TTGGGGGTTGGCGGTGGCCATGACACGCGACCGGTGGGTGGAGTGGCGTC 408

Db 332 TTGGCCCTGACCGCGGAGGTCCGAGATGGCGCCCTTGAGAGCTTGGCGGTATGGTC 273

QY 409 GATAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGACCTGGTTGAAC 468

Db 272 CTCTCGTAGTAGACAGGTGGAGCCGAAGTGGACGTAGCAGGTGCGGCGAGTGGATC 213

QY 469 CCGAATCGTATTGGAACGCTTTGTGTTTATCGCGCGGCGTTCGCGCGCAATACGGCGAC 528

Db 212 TCGAAGCGCGCGCTCGCTCGAAGCTGTAGGAGCGCAGGATCGCGGATGAGCCGC 153

QY 529 ACCGAGCGTGGATTTTCGCCGCTTGGCGGTTTCGGCGAAGCTGATCTGCTCCGCTG 588

Db 152 TTCGGAGGCGCGTACTTGGGAGTAGGTGTTCCGCGCGCGCTGAGCGCGTGCTGAC 93

QY 589 GTGGGTTTCGGCGCAGCAGCATTTGTACCGCCGCTGTCCAGCCCCCAAGGTGTGGCGTGG 648

Db 92 GCGGTGACCGCTGCGCGCGCGCGCGGCGGACGCGGAGGACGACATGCGGCGCAGCGGA 33

QY 649 ATCAACTGCTGTCGGCAGTTTGAT 674

Db 32 GACATTCCCGTTGCGGATATTGGGAT 7

RESULT 9
CD881613 720 bp mRNA linear EST 14-JUL-2003
LOCUS
DEFINITION
F1.103123F010329 F1 Triticum aestivum cDNA clone F1103L23, mRNA
sequence.

ACCESSION
CD881613
VERSION
CD881613.1 GI:32641286
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 720)

AUTHORS
TITLE
JOURNAL
COMMENT

Genoplatte.
Genoplatte, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplatte
Genoplatte

93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french
plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>)
and <http://genoplatte-info.infobiogen.fr>.

FEATURES
source

Location/Qualifiers
1..720
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="F1103L23"
/tissue_type="leaf one"
/clone_lib="F1"

ORIGIN

Query Match 6.5%; Score 46; DB 14; Length 720;
Best Local Similarity 48.1%; Pred. No. 0.54;
Matches 130; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 405 CQTGATAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGTT 464

Db 251 CQTGGAGAGAGCTGGTGGCGCTCAAGCCCAAGAGCATGACTTCGCGCAGCGCGCG 310

QY 465 GAACCCGAATGCGTATTGAGACGGTTTGTGTTATCGGCGGCGTGGCGCGCAATACGG 524

Db 311 TCTGCGCTCGCCATCGAGACCGCCCATGAGGCGCTCGAGAGCGCGGCTTCTCGCGCG 370

QY 525 CGACACCGAGCGGTGATTTTCGCGCGTGGCGGTTTCGCGCAAGCTCATCTGTTCCC 584

Db 371 CAAGTCCATCTCTGCTCGCGCGCGCGGAGTCTCGGACCTCTGTCATCCAGCTGGC 430

QY 585 GCTGTTGGTTCGCGCGCAGCATTTGTACGCCCTGTCCAGCCCCCAAGGTGTGGCG 644

Db 431 GAAGCAAGTTTACGGCGCTGCGGTTGGCGCCAGCCGCGAGCTGGAGCT 490

QY 645 CTGATCAAGCTGCTGTCGCGAGTTTGAT 674

Db 491 CQTGAAAGCTGGGAGCGCGCTGGCCAT 520

RESULT 10
CA829022/c

LOCUS
DEFINITION
1114036D07.y2 1114 - Unigene IV from Maize Genome Project Zea mays
cDNA, mRNA sequence.

ACCESSION
CA829022

VERSION
CA829022.1 GI:26457439

KEYWORDS
EST.

SOURCE
ORGANISM

Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1 (bases 1 to 620)

Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1114036 row: D column: 07.

FEATURES
source

Location/Qualifiers
1. .620
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone_lib="1114 - Unigene IV from Maize Genome Project"
/note="This library represents the unique genes found in the fourth round of EST sequencing at Stanford University for the maize genome project. Sequences are present from libraries 1091 and 3524. Contigs were assembled using ZmBAssembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."

ORIGIN

Query Match 6.2%; Score 44.4; DB 14; Length 620;
Best Local Similarity 46.0%; Pred. No. 1.4;
Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 349 TTGGCGCGTTTCGGCGTGGCCACTGACACGCGCGGTTGGGTTGAGGTGAGGTGAGGTG 408
Db 357 TTGGCGCGTTTCGGCGTGGCCACTGACACGCGCGGTTGGGTTGAGGTGAGGTGAGGTG 298
QY 409 GATAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCGCAATTCGTCGACCTGCTGTAAC 468
Db 297 CTCTCGTAGTAGACGAGTGGGAGCCGGAAGTGGACGTAGCAGGTGCGCGGAGGTGATC 238
QY 469 CGCAATCGTATTGACGCGTGGTGTGTTGTTATGCGCGGCGTGGCGCGCAATACGCGGAC 528
Db 237 TCGAAGCGCGCGTGGCGTGGTGTGTTATGCGCGGCGTGGCGCGCAATACGCGGAC 178
QY 529 ACCGAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 588
Db 177 TTCGGAGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 118
QY 589 GTGGGTTTCGGCGCAGCAGCATTTGTCACGCGCGTGTCCAGCCCGCTGTCCAGCCCGCTGG 648
Db 117 GCGGTGACCGTGGCG 58
QY 649 ATCAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 674
Db 57 GACATTGCGCGTTGCGGATATTGGAT 32

RESULT 11

BQ704105/c
LOCUS 946110C04.y1 946 - tassal primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.

ACCESSION BQ704105
VERSION BQ704105.1 GI:21843524
KEYWORDS EST.

SOURCE

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS 1 (bases 1 to 644)
TITLE Walbot, V.
JOURNAL Maize ESTs from various cDNA libraries sequenced at Stanford
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946110 row: C column: 04.
Location/Qualifiers
1. .644
/organism="Zea mays"

FEATURES
source

1. .644
/organism="Zea mays"

/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOLR"
/clone_lib="946 - tassal primordium prepared by Schmidt lab"
/note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI; Site 2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

ORIGIN

Query Match 6.2%; Score 44.4; DB 13; Length 644;
Best Local Similarity 46.0%; Pred. No. 1.4;
Matches 150; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 349 TTGGCGCGTTTCGGCGTGGCCACTGACACGCGCGGTTGGGTTGAGGTGAGGTGAGGTG 408
Db 364 TTGGCGTGGAGCGCGGAGAGGTCCGAGATGGCGCCCTTGGAGAGCTTGGCGGTTATGTC 305
QY 409 GATAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCGCAATTCGTCGACCTGCTGTAAC 468
Db 304 CTCTCGTAGTAGACGAGTGGGAGCCGGAAGTGGACGTAGCAGGTGCGCGGAGGTGATC 245
QY 469 CGCAATCGTATTGACGCGTGGTGTGTTATGCGCGCGGCGTGGCGCGCAATACGCGGAC 528
Db 244 TCGAAGCGCGCGTGGCGTGGTGTGTTATGCGCGGCGTGGCGCGCAATACGCGGAC 185
QY 529 ACCGAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 588
Db 184 TTCGGAGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 125
QY 589 GTGGGTTTCGGCGCAGCAGCATTTGTCACGCGCGTGTCCAGCCCGCTGTCCAGCCCGCTGG 648
Db 124 GCGGTGACCGTGGCG 65
QY 649 ATCAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 674
Db 64 GACATTGCGCGTTGCGGATATTGGAT 39

RESULT 12

AW927407/c
LOCUS 945001E08.Y1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.

ACCESSION AW927407
VERSION AW927407.1 GI:8102603
KEYWORDS EST.

SOURCE

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS 1 (bases 1 to 532)
TITLE Walbot, V.
JOURNAL Maize ESTs from various cDNA libraries sequenced at Stanford
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945001 row: E column: 08.
Location/Qualifiers
1. .532

```

/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH108"
/clone_lib="945" - Mixed adult tissues from Walbot lab,
same as 707 (SK)
/Note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

```

ORIGIN

```

Query Match      6.0%; Score 42.8; DB 10; Length 532;
Best Local Similarity 45.7%; Pred. No. 3.4;
Matches 149; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 349 TTGGGCGGTTGCGCGGTGGCCACTGACACGCGACCGGTGGGTTGAGGTGAGCGTC 408
    |||||
Db 332 TTGGCGCTGCACGCCGCGAGGTCCTCGAGATGGCGCTTGGAGAGCTTCCCGGTTATGGTC 273
    |||||

QY 409 GATAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGTTGAAC 468
    |||||
Db 272 CTCTCTGTAGTACAGAGTGGGAGCGAAGTGGACGTAGCAGGTGCCGCGAGTGGATC 213
    |||||

QY 469 CCGAATGCGTATTTGCGCGGTTTGTGTTATCGCGCGGTGCGCGCCCAATACGCGAC 528
    |||||
Db 212 TCGAAGCGCCCGTGGCTCGTGAAGCTGTAGGAGCGGAGTCCGGGATGAGCCCG 153
    |||||

QY 529 ACCGACGCGTGAATTTGCGCGGTGCGCGGTGGGCGGAGAGCTGATCTGTTCCGCTG 588
    |||||
Db 152 TTGGGAGGCGCTACTTGGGAGTAGTGTCTTTCGCGCGCGCTGAGCGAGCTGGCTGAC 93
    |||||

QY 589 GTGGGTTTGGCGCGAGCAGATTGTCAACCGCGCTGCCAGCCCAAGTGTGGCGCTGG 648
    |||||
Db 92 GCGGTGACGCTGGCGCGCGGCGCGCGGCGGCGGAGGACGACGATGCGGGGCGGCGA 33
    |||||

QY 649 ATCAACGCTGCTGCGCAGTTGTGAT 674
    |||||
Db 32 GACATTGCGGTTGCGGATTTGGGGT 7
    |||||

```

```

RESULT 13
CD861975
LOCUS      CD861975.1 AZ01 Triticum aestivum cDNA clone AZ01101N14,
DEFINITION mRNA sequence.
ACCESSION CD861975.1 GI:32545791
VERSION    CD861975.1
KEYWORDS   EST.
SOURCE     Triticum aestivum (bread wheat)
ORGANISM   Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Poaceae; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 538)
AUTHORS    Genoplante.
TITLE      Genoplante, a major partnership french program in plant genomics
JOURNAL    Unpublished (2003)
COMMENT    Contract: Genoplante
           Genoplante
           93, rue Henri Rochefort 91025 EVRY CEDEX France
           Tel: 33 1 69 47 54 00
           Fax: 33 1 69 47 54 10
           This sequence has been generated in the framework of the french
           plant genomics programme 'Genoplante' (http://www.genoplante.com
           and http://genoplante-info.inbio.gen.fr).
FEATURES   Location/Qualifiers
            source          1..538

```

```

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ01101N14"
/tissue_type="leaf"
/clone_lib="AZ01"

```

ORIGIN

```

Query Match      6.0%; Score 42.8; DB 14; Length 538;
Best Local Similarity 47.4%; Pred. No. 3.4; Indels 0; Gaps 0;
Matches 128; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 405 CGTGCATTAAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCAATCGTGTGCTGACTGGTT 464
    |||||
Db 108 CGTGGAGGAGAGCTGTGGCGCTCAAGCCCAAGAGCATGCACTTCGCGCAGGCGCGCG 167
    |||||

QY 465 GAACCCGATGCGTATTTGGACGCGTTTGTGTTATCGGCGGCTCGGCGCAATACGG 524
    |||||
Db 168 TCTGCCCTCGCCATCGAGAGCCGCCCATGAGGGCTCGAGAGCAGGCTTCTCGGCGCG 227
    |||||

QY 525 CGACACCGGACGCTGATTTTCGCGCTGCGCGCTTCGCGCAAGCCTGATCTGTTCCC 584
    |||||
Db 228 CAATCTCATCTTGTCTCTCGCGCGCGCGGAGTGGGACCTTGTATCCAGCTGGC 287
    |||||

QY 585 GCTGCTGGGTTTGGCGCAGCAGCATTTGTCACGCCGCTGCCAGCCCAAGGTTGGCG 644
    |||||
Db 288 AAGCAAGTTTACGCGCATCGAAGTGGCGGCCACAGCCAGCAGCCCAAGCTGGAGCT 347
    |||||

QY 645 CTGATCATGCTGCTGCGGAGCTTGTGAT 674
    |||||
Db 348 CCGAAAAGCTTGGGAGCGGACGCTCGCCAT 377
    |||||

```

RESULT 14

```

BX425603      885 bp mRNA linear EST 15-MAY-2003
BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
LOCUS      CL0BB0222A07 3-PRIME, mRNA sequence.
ACCESSION    BX425603.1 GI:30770486
VERSION      BX425603
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 885)
AUTHORS    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT    Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           Library was constructed by Life Technologies, a division of
           Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
           Paraday Avenue Genoscope sequence ID : CL0BB0222A07FPI.
FEATURES   Location/Qualifiers
            source          1..885
            organism        "Homo sapiens"
            mol_type        "mRNA"
            db_xref         "taxon:9606"
            clone            "CL0BB0222A07"
            tissue_type      "NEUROBLASTOMA"
            clone_lib        "Homo sapiens NEUROBLASTOMA"
            notes            "Vector: pCMVSPORT 6; 1st strand cDNA was primed
            with a NotI-oligo(dT) primer. Five prime end enriched
            double-strand cDNA was digested with Not I and cloned into
            the Not I and EcoRV sites of the pCMVSPORT 6 vector.
            Library was not normalized."

```

ORIGIN

```

source      1. .910  

            /organism="Drosophila melanogaster"  

            /mol_type="genomic DNA"  

            /db_xref="taxon:7227"  

            /clone="BACR14J21"  

            /clone_lib="RPCI-98"  

            /note="end : T7"  

  

ORIGIN  

Query Match      5.8%; Score 41.2; DB 29; Length 910;  

Best Local Similarity 14.8%; Pred.No.11;  

Matches 47; Conservative 150; Mismatches 119; Indels 2; Gaps 1;  

  

QY      333 GCCCATGACACACCCTTTGGCGGTTCGCCGTGCACATGACAGCGCAACCCGGTGGC 392  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

Db      518 GGSSSKGGCGSGSYGGGGGGTGTSGGGGGGSSSSGCSSSSSSSSSSCGSSBGCB 577  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

QY      393 GTGTGAGGTGAGCGTTCGATAAGACAGCGGGTTTGGGTAAACCCCATTGTGTATGCGAATCGT 452  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

Db      578 SSCGCGCSVSYSSTGTS GCSCGSSSSSSSSSSS SCSSTSSTSGCGSSGSHSSC 637  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

QY      453 GCTGACC--TGTTGAACCCGAATCGGTATTGACAGCGTTTGTGTTATCGCGCGCGTC 510  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

Db      638 SSTBTSSSTTETBKSTSYSTSYSTSTSTSTSTSBFSYGSBHKSTSTSSSTSSSSTSS 697  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

QY      511 GGCGCGCAATACGGCGACACCGAGCGGTGGATTTCGCCGTGCGCGCTCGCGCAAGC 570  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

Db      698 SSSSTSSGCTSTBYSTSTBSYSBTBTBTSTSTBTBTSTBYTBTTBYTKSTTSMST 757  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

QY      571 CTGATCTGTTCCCGCTGGTGGTTTCGGCGCAGCAGCATTTGACGCCGCTGCACG 630  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

Db      758 YTTBSTSKBSTBTAKSTGKNTSBTTCTSSSSBSSTSYSTSCBSBSBSSTSY 817  

        |||::||::||::||::||::||::||::||::||::||::||::||:  

  

QY      631 CCCAAGGTGTGGCGCTGG 648  

        |||::||::||::||::||  

Db      818 SBCTSTSTSTSSBSBS 835  

        |||::||::||::||::||

```

Search completed: March 16, 2004, 02:52:15
Job time : 1958.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:15:20 ; Search time 3495.31 Seconds
(without alignments)
10825.501 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1421_2293

Perfect score: 873
Sequence: 1 gtccataagcagcggttg.....ctgtgagctctgaccgtag 873

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_ey:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	873	100.0	2374	1	CGLYSEG	X96471 C.glutamici
2	873	100.0	2374	6	AG9393	A93933 Sequence 2
3	873	100.0	333150	1	AP005277	AP005277 Coryneb
4	873	100.0	349980	6	AX127147	AX127147 Sequence
5	568	65.1	627	6	AX123538	AX123538 Sequence
6	568	65.1	627	6	BD165655	BD165655 Novel pol
7	568	65.1	1083	6	AX813971	AX813971 Sequence
8	404.2	46.3	993	6	AX067087	AX067087 Sequence
9	329.2	37.7	308650	1	AP005218	AP005218 Coryneb
10	326	37.3	822	6	AX063771	AX063771 Sequence
11	326	37.3	822	6	AX244059	AX244059 Sequence
12	306	35.1	711	6	AX643028	AX643028 Sequence
13	304.4	34.9	712	6	AX643030	AX643030 Sequence
14	303	34.7	708	6	AX123539	AX123539 Sequence
15	303	34.7	708	6	BD165656	BD165656 Novel pol
16	294.6	33.7	349535	1	EX248357	EX248357 Coryneb
17	193.6	22.2	1568	6	E54483	E54483 Heat-resist
18	193.6	22.2	1771	1	AB083133	AB083133 Coryneb
19	141.2	16.2	345783	1	AP003001	AP003001 Mesorhizo
20	136.6	15.6	8472	1	AE008076	AE008076 Agrobacte
21	136.6	15.6	13051	1	AE009111	AE009111 Agrobacte
22	133.6	15.3	302007	1	SC039132	AL939132 Streptomy
23	115	13.2	306250	1	SMES91788	AL591788 Sinorhizo
24	89.6	10.3	298300	1	AP005025	AP005025 Streptomy
25	88.6	10.1	10871	1	AE011965	AE011965 Xanthomon
26	88.4	10.1	303550	1	SC039131	AL939131 Streptomy
27	87.2	10.0	1525	1	PSE1GRA	AE002063 Deinococc
28	86.2	9.9	12257	1	AE002063	AE005530 Escherich
29	84.8	9.7	10336	1	AE005530	AE015317 Shigella
30	83.6	9.6	10278	1	AE015317	AP002563 Escherich
31	83.6	9.6	266658	1	AP002563	AE016988 Shigella
32	83.6	9.6	292906	1	AE016988	AE016766 Escherich
33	82	9.4	303121	1	AE016766	AE000382 Escherich
34	80.4	9.2	11024	1	AE000382	U28377 Escherichia
35	80.4	9.2	141744	1	ECU28377	AC020884 Mus muscu
36	79.4	9.1	110000	2	ACU20884_0	AE016857 Pseudomon
37	78.4	9.0	334108	1	AE016857	AL627274 Salmonell
38	76.4	8.8	256050	1	AL627274	AE016835 Salmonell
39	76.4	8.8	300431	1	AE016835	AE008808 Salmonell
40	74.8	8.6	22929	1	AE008808	AL646059 Ralstonia
41	73.4	8.4	190050	1	AL646059	AE016937 Bacteroid
42	73.4	8.4	305961	1	AE016937	AR012071 Sequence
43	71.6	8.2	5541	6	AR012071	AR025196 Sequence
44	71.6	8.2	5541	6	AR025196	AR038653 Sequence
45	71.6	8.2	5541	6	AR038653	

ALIGNMENTS

RESULT 1

CGLYSEG 2374 bp DNA linear BCT 19-MAR-2001
LOCUS C.glutamici lysE and lysG genes.
DEFINITION X96471
ACCESSION X96471.1 GI:1729753
VERSION lysE gene, lysG gene; Lysine export regulator protein; Lysine
KEYWORDS exporter protein; Lysine governor.
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Vrljic, M., Sahm, H. and Eggeling, L.
TITLE A new type of transporter with a new type of cellular function:

JOURNAL	L-lysine export from <i>Corynebacterium glutamicum</i>	
MEDLINE	Mol. Microbiol. 22 (5), 815-826 (1996)	
PUBMED	897126810	
REFERENCE	89711704	
AUTHORS	2 (bases 1 to 2374)	
TITLE	Direct Submission	
JOURNAL	Submitted (07-Mar-1996) M.M. Vrljic, Institut fuer Biotechnologie	
FEATURES	1, Forschungszentrum Juelich, Postfach 1913, D-52425 Juelich, FRG	
source	Location/Qualifiers	
	1..2374	
	/organism="Corynebacterium glutamicum"	
	/mol_type="genomic DNA"	
	/strain="R127"	
gene	/db_xref="taxon:1718"	
	complement (82..954)	
CDS	/gene="lysG"	
	complement (82..954)	
	/gene="lysG"	
	/note="lysine governor"	
	/codon_start=1	
	/transl_table=11	
	/product="lysine export regulator protein"	
	/protein_id="CAA65323.1"	
	/db_xref="GI:1729754"	
	/db_xref="GOA:P94632"	
	/db_xref="SWISS-PROT:P94632"	
	/translation="MNPIQLDILSIIDEGSFEGLASLISPSAVSRVKALEHHVG RVLVSRTQPAKAEATVLEQVAKMVLQAEAKQLSGRLAEPLITIAINADLSLTW RFPVNEVAGSGATLRLDEAHTLSLRGDLVIGAVTREAHPVAGGVVVELGTWR HLAIVTSLRDYVMDGLDWAAMPVLRFGKVDLQDLRDLDGRVDGPVGRVRSIVPS AFGEALREGLGWLPLETQAPMLKAGEVILLDEIIDPTPMYQWRKLESRLARL TDAVDAIAEGLSP"	
gene	1025..1726	
	/gene="lysE"	
CDS	1025..1726	
	/gene="lysE"	
	/codon_start=1	
	/transl_table=11	
	/product="lysine exporter protein"	
	/protein_id="CAA65324.2"	
	/db_xref="GI:13397387"	
	/db_xref="GOA:P94633"	
	/db_xref="SWISS-PROT:P94633"	
	/translation="MEIFITGLIGASLISIGPNVLVIKQIKRGLIIVLVCLL SVDFIFAGTLGVDLISNAPIVLDINRWGIIAYLLFNFAWAAKDANKVAPOLIE ETEPIDPDTPLGSSAVATDTRNRVEVSDVKQVWVPMALVILTNLNPALIDA FVFIGGVAQYGDTRGWIIPAGAFASLIWFLVLFVFGAALLSLPSLPKVRNINVVV AVVMTALAIIKLMLNG"	
gene	complement (1723..2352)	
CDS	complement (1723..2352)	
	/gene="orf3"	
	complement (1723..2352)	
	/gene="orf3"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="CAA65325.1"	
	/db_xref="GI:1729756"	
	/db_xref="SPTFEMBL:P94634"	
	/translation="MYALRDIYVASKALYVGISSYGPETLTAEEAFMAEGCPILIHQ PYSIINWEEPGDDGEMLOSAANGLGVIAFPLAQGLITDKYLDGIPGSSRASQ GKSLSEGMLNINMVKINDIAQERQSLAQMALAWYLRQEGYGDVTVTSALIGA SSVEQLDLSLNLNLEFSDAELEAIDRSHDAGININWAKATDSTREN"	
ORIGIN		
Query Match	100.0%; Score 873; DB 1; Length 2374;	
Best Local Similarity	100.0%; Pred. No. 9.8e-185;	
Matches 873; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 GTCCATAGCAGCGGTTGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGGTTG 60	
Db	1421 GTCCATAGCAGCGGTTGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGGTTG 1480	
Qy	61 AACCCGAATCGCTATTGGACGCGCTTTGTTGTTATCGCGCGCGCTCGGCGCAATACGGC 120	
Db	1481 AACCCGAATCGCTATTGGACGCGCTTTGTTGTTATCGCGCGCGCTCGGCGCAATACGGC 1540	
Qy	121 GACACCGACGGTGGATTTTCGCCGCTGCGCGCTTCGCGCAAGCTCATCTGTTCCCG 180	
Db	1541 GACACCGACGGTGGATTTTCGCCGCTGCGCGCTTCGCGCAAGCTCATCTGTTCCCG 1600	
Qy	181 CTGTTGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTTCAGCCCCCAAGGTTGGCGC 240	
Db	1601 CTGTTGGTTTCGGCGCAGCAGCATTTGTCAGCCCGCTTCAGCCCCCAAGGTTGGCGC 1660	
Qy	241 TGGATCAACGCTCGTGGCGAGTTTGATGATCGCGCATTTGCCCATCAAACTGATCTTCATG 300	
Db	1661 TGGATCAACGCTCGTGGCGAGTTTGATGATCGCGCATTTGCCCATCAAACTGATCTTCATG 1720	
Qy	301 GGTAGTTTCGCGGTTTGGAAATCGTGGCTTCGCCCAAAATGTTGATSCCGCGGTG 360	
Db	1721 GGTAGTTTCGCGGTTTGGAAATCGTGGCTTCGCCCAAAATGTTGATSCCGCGGTG 1780	
Qy	361 TGGGAATCTCATCGATCGCTCAACTCGCGCTCAGAAACTCCAAAGTTGTTGAGTGAA 420	
Db	1781 TGGGAATCTCATCGATCGCTCAACTCGCGCTCAGAAACTCCAAAGTTGTTGAGTGAA 1840	
Qy	421 TCAAGGCTGTTTCAGCTGCTCAACTGACGAGCACCAATCAATGCACTGGTCAOAGTA 480	
Db	1841 TCAAGGCTGTTTCAGCTGCTCAACTGACGAGCACCAATCAATGCACTGGTCAOAGTA 1900	
Qy	481 TCCGCGCGTACTCTCTTGTGCGCAGCAGCCCATGCAAGCGCCCATCTGCGGAAGTGAC 540	
Db	1901 TCCGCGCGTACTCTCTTGTGCGCAGCAGCCCATGCAAGCGCCCATCTGCGGAAGTGAC 1960	
Qy	541 TGCCCGGTTTCCTGGCGATGTCATTTAGCTTCGCGACCATATCAATATTGTTCACTTC 600	
Db	1961 TGCCCGGTTTCCTGGCGATGTCATTTAGCTTCGCGACCATATCAATATTGTTCACTTC 2020	
Qy	601 AACATGCTCTCAGACAGGACTTACCTTGGTGGCGGGGAAACCTCTGGGAATTCATTCG 660	
Db	2021 AACATGCTCTCAGACAGGACTTACCTTGGTGGCGGGGAAACCTCTGGGAATTCATTCG 2080	
Qy	661 AGATATTGTCGTGACGAGCCCTGCGCAAGTGTGAGAAAGCAATGACGCAAGACCA 720	
Db	2081 AGATATTGTCGTGACGAGCCCTGCGCAAGTGTGAGAAAGCAATGACGCAAGACCA 2140	
Qy	721 TTGTTGGAGCTGATGCAACAGTTCTCACGCTCATCGCGCTTCCTCCACCCACGA 780	
Db	2141 TTGTTGGAGCTGATGCAACAGTTCTCACGCTCATCGCGCTTCCTCCACCCACGA 2200	
Qy	781 TTAATGATGAATAGCTTGGCTGATGATCAGAGCGGCGAGCCCTCTCCGCCATGAAC 840	
Db	2201 TTAATGATGAATAGCTTGGCTGATGATCAGAGCGGCGAGCCCTCTCCGCCATGAAC 2260	
Qy	841 TCAGCGCCCTCCGCTGTGAGCTCTGACCGTAG 873	
Db	2261 TCAGCGCCCTCCGCTGTGAGCTCTGACCGTAG 2293	
RESULT 2		
A93933		
LOCUS	Sequence 2 from Patent WO973597.	
DEFINITION	2374 bp DNA linear PAT 22-JAN-2000	
ACCESSION	A93933	
VERSION	A93933.1 GI:6742037	
KEYWORDS		
SOURCE	<i>Corynebacterium glutamicum</i>	
ORGANISM	<i>Corynebacterium glutamicum</i>	
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
	<i>Corynebacterineae</i> ; <i>Corynebacteriaceae</i> ; <i>Corynebacterium</i> .	
REFERENCE	1 (bases 1 to 2374)	
AUTHORS	Vrljic, M. and Eggeling, L.	
TITLE	PROCESS FOR THE MICROBIAL PRODUCTION OF AMINO ACIDS BY BOOSTED	
JOURNAL	ACTIVITY OF EXPORT CARRIERS	
	Patent: WO 9723597-A 2 03-JUL-1997;	
	KERNFORSCHUNGSANLAGE JUELICH (DE); VRLIJC MARINA (DE)	

FEATURES		Location/Qualifiers	
source		1..2374	
ORIGIN		/organism="Corynebacterium glutamicum" /mol_type="unassigned DNA" /db_xref="taxon:1718"	
Query Match		100.0%; Score 873; DB 6; Length 2374;	
Best Local Similarity		100.0%; Pred. No. 9.8e-185;	
Matches 873; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GTCCGATAGCAGCGGTTGGGTAAAGCCATGTTGATGGCAATCGTGTGACCTGGTGG 60	
DB	1421	GTCCGATAGCAGCGGTTGGGTAAAGCCATGTTGATGGCAATCGTGTGACCTGGTGG 1480	
QY	61	AACCCGAATCGTATTGGACGCGTTTGTGTTATCGCGCGCTCGCGCGCAATACGCGC 120	
DB	1481	AACCCGAATCGTATTGGACGCGTTTGTGTTATCGCGCGCTCGCGCGCAATACGCGC 1540	
QY	121	GACACCGAGCGGTGGATTTTCGCGCTGGCGGTTTCGCGGCAAGCTGATCTGGTCCCG 180	
DB	1541	GACACCGAGCGGTGGATTTTCGCGCTGGCGGTTTCGCGGCAAGCTGATCTGGTCCCG 1600	
QY	181	CTGGTGGGTTTCGCGCGACGACATTTGTACGCGCGCTGTCCAGCCCGCTGTCAGCGC 240	
DB	1601	CTGGTGGGTTTCGCGCGACGACATTTGTACGCGCGCTGTCCAGCCCGCTGTCAGCGC 1660	
QY	241	TGGATCAACGCTGCTGGCGAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300	
DB	1661	TGGATCAACGCTGCTGGCGAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 1720	
QY	301	GGTAGTATTTTCGCGGGTTTGGATCGGTGGCTTCGCGGCAAGTTGATGCGCGGCTCG 360	
DB	1721	GGTAGTATTTTCGCGGGTTTGGATCGGTGGCTTCGCGGCAAGTTGATGCGCGGCTCG 1780	
QY	361	TGGGAATCTCATCGATCGCTCCAACTCGCGCTCAGAACTCCAAAGTTGTCAGTGAA 420	
DB	1781	TGGGAATCTCATCGATCGCTCCAACTCGCGCTCAGAACTCCAAAGTTGTCAGTGAA 1840	
QY	421	TCAAGGCTGTTGTCAGCTGCTCACTGACGAGCAAGCAATCAATGCACTGGTCACGGTA 480	
DB	1841	TCAAGGCTGTTGTCAGCTGCTCACTGACGAGCAAGCAATCAATGCACTGGTCACGGTA 1900	
QY	481	TCGCGCGCTGATCTCTCTGCTCGCGACACCCATGCAAGCGCCATCTGGCGAAGTGAC 540	
DB	1901	TCGCGCGCTGATCTCTCTGCTCGCGACACCCATGCAAGCGCCATCTGGCGAAGTGAC 1960	
QY	541	TGCGCGGCTTCTCGCGCGATGTCATTCAGCTTGGGACCATATCAATATTTGTCACGTTTC 600	
DB	1961	TGCGCGGCTTCTCGCGCGATGTCATTCAGCTTGGGACCATATCAATATTTGTCACGTTTC 2020	
QY	601	AACATGCTCTCAGACAGGACTTACCTCGCTGGCGGGGAACCTCTCGAATTCATCG 660	
DB	2021	AACATGCTCTCAGACAGGACTTACCTCGCTGGCGGGGAACCTCTCGAATTCATCG 2080	
QY	661	AGATATTGTCGTCAGCAGCCCTCGCAAGTGTGAGAAACATGACCCAGACCA 720	
DB	2081	AGATATTGTCGTCAGCAGCCCTCGCAAGTGTGAGAAACATGACCCAGACCA 2140	
QY	721	TTGTTGGCAGCTGATCTCAACAGTTTCTCAACGCTCATCGCGCGGTTCTCCACCCACCA 780	
DB	2141	TTGTTGGCAGCTGATCTCAACAGTTTCTCAACGCTCATCGCGCGGTTCTCCACCCACCA 2200	
QY	781	TTAATGATGGAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCCGCCATGAC 840	
DB	2201	TTAATGATGGAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCCGCCATGAC 2260	
QY	841	TCAGCGCTCTCGCTGTGAGCTCTGACCGTAG 873	
DB	2261	TCAGCGCTCTCGCTGTGAGCTCTGACCGTAG 2293	

(IISF) Family protein
 TIGR00968:3a0106e01: sulfate transport system permease protein
 TIGR01166:chiO: cobalt transport protein ATP-binding subunit
 TIGR01184:ntrCD: nitrate transport ATP-binding subunits C and D
 TIGR01186:proV: glycine betaine/L-proline transport ATP binding subunit
 TIGR01187:potA: spermidine/putrescine ABC transporter ATP-binding subunit
 TIGR01188:draA: daunorubicin resistance ABC transporter ATP-binding subunit
 TIGR01189:cmA: heme exporter protein CcmA
 /codon_start=1
 /transl_table=11
 /product="ABC-type transporter, ATPase component"
 /protein_id="BAB98339.1"
 /db_xref="GI:21323713"
 /translation="MTTTPADVTDLRTVGYTAVKGLNFHVQGEVFGLLGTGAG KTSLEVEGLSPSGTVRISGLDPVADRAILRELGIMLOSGLPSQLVASTMDM WGTCTTPRAIKDVLADVLLHRENVKVGALSGGQRRLDLCALLGDPSTILFDEFT TGLDPESETRTQQLLDLQKQVMTLTTTHLEAEFLCDRIAMNAGIAVEGLTDE LVAREKIIISFVLRGQVLEPLVSGAIIIRDNNHRIATTTLQHTLELTAAETGAI ALEGFAKPAATLESVFMDIASLENTSLQTA"
 4003..4836
 /gene="Cg10947"
 4003..4836
 /gene="Cg10947"
 /codon_start=1
 /transl_table=11
 /product="ABC-type transporter, permease components"
 /protein_id="BAB98340.1"
 /db_xref="GI:21323714"
 /translation="MTTSHARGLOHATPERKTSFFKTSIFAELWQPRNKULLPM LTFPVGPIILLILGNGGAESANSDFYVMTLLFQVYTLVSWATRRDRLVKR AVTGAEADIDIIIGACIFGALLITFTVPIILMLVAGAPINLVPIVFAVLGLLL CSALALGTGFTNRAEAQMTSMFVLMGLGSRFVFGDSIVADILATVPPAAIS DLVQIGWAGATFADSVGGVEAANFAGIFQDMLPIGLIGLAATAAVAAANRVMWDSY R"
 4879..6138
 /gene="Cg10948"
 4879..6138
 /gene="Cg10948"
 /EC_number="2.7.3.-"
 /codon_start=1
 /transl_table=11
 /product="Two-component system, sensory transduction histidine kinases"
 /protein_id="BAB98341.1"
 /db_xref="GI:21323715"
 /translation="MNKDFWTAGTARWFSRGVSLASPVTPAPLNSWRRLPNLYTL YTRVSLQAIPLVLLSYFLGIVANAGTLPNSFVLLGPSVILLITVLVVEYQPSLNS HPRRSVQFFTGLVLNLGVSVVQLIPGLMNSDNRATLFLITCYFLLSYAVI PMNRYVNLIASVAVLMTSTTDYLSALVWVPIPLMAGTVRLSVTVDMVKEVERK RELEASELVEERLPAQELHDITLQHLAAMSVKSELALAKRGDDRLNLELQK LRTSMSEMDRVGVYRTVNLATIEGAKSLADAHILHSVIGTTSQVSPAHRELCAW LVREATTNHLRSDATDALTLSSTEVMDNNGVKNKIDIGRLSGLSALRSRBSAGMTL IYSREDQPSVRMLINAPANTPAKEA"
 6139..6750
 /gene="Cg10949"
 6139..6750
 /gene="Cg10949"
 /note="PF00072:Response regulator receiver domain
 PF00196:Bacterial regulatory proteins, luxR family"
 /codon_start=1
 /transl_table=11
 /product="Two-component system, response regulators consisting of a CheY-like receiver domain and a HTH DNA-binding domain"
 /protein_id="BAB98342.1"
 /db_xref="GI:21323716"
 /translation="MTSISIADEALIASLIATLLSLBPDLDVRPTAGSGEELIETWA

DPNSRTDVCVLDLQGGIDGTDATRLMETTPDPLAVLITVSHARPQLKRALAGVLG FLPTKSTADFAITATVHAGRYVIDPELAAMTISAGESPLTNREBEVLELAGQSLA REIAVAHAHAPGTRNYLSQAMTKVGQNRFEAFTRARELGL"
 complement(6781..6987)
 /gene="Cg10950"
 complement(6781..6987)
 /gene="Cg10950"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB98343.1"
 /db_xref="GI:21323717"
 /translation="MKHYSMYFGAGPLANTPHTEKVFMPDIVSIVNHFENWSTNNWVN MTGLPRTFPVAFVSGSTVLAVLSF"
 7048..8205
 /gene="Cg10951"
 7048..8205
 /gene="Cg10951"
 /codon_start=1
 /transl_table=11
 /product="rRNA-processing ribonuclease BN"
 /protein_id="BAB98344.1"
 /db_xref="GI:21323718"
 /translation="MTAMLCSPSGVQDEYDMNGVQPOEHLDAITLADPHGNPEN SGRKERLNFQGHKALNRTVDFPDGLDLAALLTFISLTAPAVLLGTSITIF LASDSTELNLNDEVNQYVPEDQSHVNGVIDSIAGSAAAGQGVAVGVITALTSS AYVRAFRCANAVGRSBRGLTKRWMLFLNLALLGIIITLVSWNLVMTLVMGIF APIAPLHLNVLNPLTDREMPITWVRPPIVGLIMFVATLYWAPNRPKPRML SLSGFLAVGLLGLNLFYETLFAFSSYAGVGLLAVFIALWVFNLCIIIGLKID VEISRAKQLQGMQPAEDISLVPRPSIEKAKKQKQKQRLMDQAAARESEN"
 8672..10114
 /gene="Cg10952"
 8672..10114
 /gene="Cg10952"
 /note="PF00005:ABC transporter
 PF00664:ABC transporter transmembrane region."
 /codon_start=1
 /transl_table=11
 /product="ABC-type multidrug/protein/lipid transport system, ATPase component"
 /protein_id="BAB98345.1"
 /db_xref="GI:21323719"
 /translation="MKRTHITASSILRMRORGVAFGAFGLGMQLSEALVPIAL GLVDHAVLTKDLRLVGLVAVLVFVLSFSGFRGALNRANRVNFSALRUEVAD HALKNDLRNLVFGVMSRSTADASTRIQIGITGVSAATGLAATVLLISDLV LGKLLVLPVITSGVVALASKISRSVTQOEKLASGASQDMMGLRVIKAIIGGERW AVKTEKASQASARAADVAVASGVAGIGELSIANVLAALLGMRVTTGELGPGQ LIAIVGAVLISPHRIILSNISASAIHGAERVANFLNDESOAVESSETINDGE FLVIVPPASTLPBGDNLATPHADIEFTLRNSIMNHEDNVPIDPQVIRASGLTDI IVDGLDAPVRDTGSLSGQRQVRLARALHADAEVLVLMPTSDVSDVSTVSTIAQG IKQLRAGKTTIVVSSSPAFYNIADRVISHV"
 10107..11843
 /gene="Cg10953"
 10107..11843
 /gene="Cg10953"
 Query Match 100.0%; Score 873; DB 1; Length 333150;
 Best Local Similarity 100.0%; Pred. No. 7,7e-185;
 Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GTCCGTAAGCGCGGTTTGGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGTTG 60
 Db 328598 GTCCGTAAGCGCGGTTTGGGTAAAGCCCATGTTGATGCGCAATCGTCTGACCTGTTG 328599
 Qy 61 AACCCGAATCGCTATTTCGACGCGGTTTGTGTTTATCGCGCGCGCTCGCGCGCAATACGC 120
 Db 328598 AACCCGAATCGCTATTTCGACGCGGTTTGTGTTTATCGCGCGCGCTCGCGCGCAATACGC 328479
 Qy 121 GACACCGACGCTGATTTTCGCGCGCTTCGCGCGGAGCCTGATCTGCTCCCG 180
 Db 328478 GACACCGACGCTGATTTTCGCGCGCTTCGCGCGGAGCCTGATCTGCTCCCG 328419
 Qy 181 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGCTGTCCAGCCCCCAAGGTGTGCGCG 240

Db 328418 CTGGTGGTTTCGGGCGAGCAGCATTTGTACCGCCGCTGTCCAGCCCAAGGTGTGGCG 328359
Qy 241 TGGATCAACGTGTGTGGGCGATGTGATGACCGCATGTGCAATCAAACTGATGTGATG 300
Db 328358 TGGATCAACGTGTGTGGGCGATGTGATGACCGCATGTGCAATCAAACTGATGTGATG 328299
Qy 301 GGTAGTATTTTCGGGGTTTTCGAAATCGGTGGCTTCGCCCAAAATGTGATGCGCGCTCG 360
Db 328298 GGTAGTATTTTCGGGGTTTTCGAAATCGGTGGCTTCGCCCAAAATGTGATGCGCGCTCG 328239
Qy 361 TGGGAAATCTCATGCAATCGCTCCAACTCGGGGTGAGAAATCTCCAAAGTTGTGAGTGA 420
Db 328238 TGGGAAATCTCATGCAATCGCTCCAACTCGGGGTGAGAAATCTCCAAAGTTGTGAGTGA 328179
Qy 421 TCAAGGCTGTGTGCAAGTGTCTCAATCTGAGGAGAGACCAATCAATGCACTGTGACGGTA 480
Db 328178 TCAAGGCTGTGTGCAAGTGTCTCAATCTGAGGAGAGACCAATCAATGCACTGTGACGGTA 328119
Qy 481 TCCGCGCGTACTCTCTTGTCTGCGCAGCAGCCCAATGCAAGCGCAATCTGCGCAAGTGAC 540
Db 328118 TCCGCGCGTACTCTCTTGTCTGCGCAGCAGCCCAATGCAAGCGCAATCTGCGCAAGTGAC 328059
Qy 541 TCCGCGCGTACTCTCTTGTCTGCGCAGCAGCCCAATGCAAGCGCAATCTGCGCAAGTGAC 600
Db 328058 TCCGCGCGTACTCTCTTGTCTGCGCAGCAGCCCAATGCAAGCGCAATCTGCGCAAGTGAC 327999
Qy 601 AACATGCTGTGCAAGGAGCTTACCTGTGCGGCGGAGAGCCCTCTGGAATTCATGTCG 660
Db 327998 AACATGCTGTGCAAGGAGCTTACCTGTGCGGCGGAGAGCCCTCTGGAATTCATGTCG 327939
Qy 661 AGATATTTGTCCGTGAGAGGCTTACCTGTGCGGCGGAGAGCCCTCTGGAATTCATGTCG 720
Db 327938 AGATATTTGTCCGTGAGAGGCTTACCTGTGCGGCGGAGAGCCCTCTGGAATTCATGTCG 327879
Qy 721 TTGTGGAGCTGACTGCAACAAAGTTCTCACCGTCAATCGCCGGTTCTCCACCAACGA 780
Db 327878 TTGTGGAGCTGACTGCAACAAAGTTCTCACCGTCAATCGCCGGTTCTCCACCAACGA 327819
Qy 781 TTATGATGGAATAGCTGGCTGATGATCAAGAGGAGCCCTCTCCGCGCATGAC 840
Db 327818 TTATGATGGAATAGCTGGCTGATGATCAAGAGGAGCCCTCTCCGCGCATGAC 327759
Qy 841 TCAGCGCGCTCCGCTGTGAGCTCTGAGCCGTAG 873
Db 327758 TCAGCGCGCTCCGCTGTGAGCTCTGAGCCGTAG 327726

RESULT 4
AX127147/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3,309,400) split in 11, seq 7063
1.200.001 1.549.980"

AX127147 349980 bp DNA linear PAT 11-MAY-2001
Sequence 7063 from Patent EP1108790.
AX127147 AX114121
AX127147.1 GI:14041135
Corynebacterium glutamicum
Corynebacterium Glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
1
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
Novel polynucleotides
Patent: EP 1108790-A 7063 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
Location/Qualifiers
1. 349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3,309,400) split in 11, seq 7063
1.200.001 1.549.980"

Query Match 100.0%; Score 873; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 7.7e-185; Indels 0; Gaps 0;
Matches 873; Conservative 0; Mismatches 0; Gaps 0;
Qy 1 GTGCATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTG 60
Db 128548 GTGCATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGACCTGGTTG 128489
Qy 61 AACCCGAATGCGTATTGGAAGCGCTTTGTTATCGGGGGCGTCCGGCGCAATAACGCG 120
Db 128488 AACCCGAATGCGTATTGGAAGCGCTTTGTTATCGGGGGCGTCCGGCGCAATAACGCG 128429
Qy 121 GACACCGAGCGTGGATTTTTCGCGCTGCGCGCTTTCGCGCAAGCCTGATCTGTTCCCG 180
Db 128428 GACACCGAGCGTGGATTTTTCGCGCTGCGCGCTTTCGCGCAAGCCTGATCTGTTCCCG 128369
Qy 181 CTGTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCTGTTCAGCCCCCAAGGTGTGGCG 240
Db 128368 CTGTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCTGTTCAGCCCCCAAGGTGTGGCG 128309
Qy 241 TGGATCAACGTGTGCTGTCGAGTGTGATGACCGCATTTGGCCCATCAAACTGATGTGATG 300
Db 128308 TGGATCAACGTGTGCTGTCGAGTGTGATGACCGCATTTGGCCCATCAAACTGATGTGATG 128249
Qy 301 GGTAGTATTTTCGGGGTTTTCGAAATCGGTGGCTTCGCCCAAAATGTGATCGCGCGTCG 360
Db 128248 GGTAGTATTTTCGGGGTTTTCGAAATCGGTGGCTTCGCCCAAAATGTGATCGCGCGTCG 128189
Qy 361 TGGGAAATCTCATGCAATCGCTCCAACTCGGGGTGAGAAATCTCAAGTTGTGAGTGA 420
Db 128188 TGGGAAATCTCATGCAATCGCTCCAACTCGGGGTGAGAAATCTCAAGTTGTGAGTGA 128129
Qy 421 TCAAGGCTGTGTGCAAGGAGCTTACCTGTGCGGCGGAGAGCCCTCTGGAATTCATGTCG 480
Db 128128 TCAAGGCTGTGTGCAAGGAGCTTACCTGTGCGGCGGAGAGCCCTCTGGAATTCATGTCG 128069
Qy 481 TCCGCGCGTACTCTCTTGTCTGCGCAGCAGCCCAATGCAAGCGCCATCTCCGCAAGTGAC 540
Db 128068 TCCGCGCGTACTCTCTTGTCTGCGCAGCAGCCCAATGCAAGCGCCATCTCCGCAAGTGAC 128009
Qy 541 TCCGCGCGTACTCTCTTGTCTGCGCAGCAGCCCAATGCAAGCGCCATCTCCGCAAGTGAC 600
Db 128008 TCCGCGCGTACTCTCTTGTCTGCGCAGCAGCCCAATGCAAGCGCCATCTCCGCAAGTGAC 127949
Qy 601 AACATGCTGTGCAAGGAGCTTACCTGTGCGGCGGAGAGCCCTCTGGAATTCATGTCG 660
Db 127948 AACATGCTGTGCAAGGAGCTTACCTGTGCGGCGGAGAGCCCTCTGGAATTCATGTCG 127889
Qy 661 AGATATTTTCGCGCGCAGCGCTTCGCGCAAGTGTGAGAAAGCAATGACGCCAAGACA 720
Db 127888 AGATATTTTCGCGCGCAGCGCTTCGCGCAAGTGTGAGAAAGCAATGACGCCAAGACA 127829
Qy 721 TTGTGGGAGCTGACTGCAACAAAGTTCTCACCGTCAATCGCCGGTTCTCCACCAACGA 780
Db 127828 TTGTGGGAGCTGACTGCAACAAAGTTCTCACCGTCAATCGCCGGTTCTCCACCAACGA 127769
Qy 781 TTATGATGGAATAGCTGGCTGATGATCAAGAGGAGCCCTCTCCGCGCATGAC 840
Db 127768 TTATGATGGAATAGCTGGCTGATGATCAAGAGGAGCCCTCTCCGCGCATGAC 127709
Qy 841 TCAGCGCGCTCCGCTGTGAGCTCTGAGCCGTAG 873
Db 127708 TCAGCGCGCTCCGCTGTGAGCTCTGAGCCGTAG 127676

RESULT 5
AX123538/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

AX123538 627 bp DNA linear PAT 11-MAY-2001
Sequence 3454 from Patent EP1108790.
AX123538
AX123538.1 GI:14041026

SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

REFERENCE
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotides
 JOURNAL Patent: EP 1108790-A 3454 20-JUN-2001;
 KYOMA HAKKO KOGYO CO., LTD. (JP)

FEATURES
 Location/Qualifiers
 source 1..627
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"

ORIGIN
 Query Match 65.1%; Score 568; DB 6; Length 627;
 Best Local Similarity 100.0%; Pred. No. 1.7e-116;
 Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

306 GTTTTCGCGGGTTTGAATCGGTGCGCTTCGCCCAAAATGTTGATGCGCGCGTGTGGGA 365
 Db GTTTTCGCGGGTTTGAATCGGTGCGCTTCGCCCAAAATGTTGATGCGCGCGTGTGGGA 568

366 AATCTCATCGATCGCTCCAACTCGCGGTGAGAAACTCCAAAGTTGTTGAGTGAATCAAG 425
 Db AATCTCATCGATCGCTCCAACTCGCGGTGAGAAACTCCAAAGTTGTTGAGTGAATCAAG 508

426 GCTGTTGTCAGCTGCTCAACTGACGAAGCACCATCAATGACACTGCTGACGGTATCCGC 485
 Db GCTGTTGTCAGCTGCTCAACTGACGAAGCACCATCAATGACACTGCTGACGGTATCCGC 448

486 GCGGTACTCTCTCTGTCGCGAGCACCATCGAAGCGCATCTGCGCAAGTGAATCAAT 545
 Db GCGGTACTCTCTCTGTCGCGAGCACCATCGAAGCGCATCTGCGCAAGTGAATCAAT 388

546 GCGTTCGCGGAGTCAATGAGCTGCGGAGCACCATCAATGAGTCAATGTTGACGTTCAACAT 605
 Db GCGTTCGCGGAGTCAATGAGCTGCGGAGCACCATCAATGAGTCAATGTTGACGTTCAACAT 328

606 GCCCTCAGACAGGAGCTTACCTGCGCGCGGAGAACCTCTGGAATTCATCGAGATA 665
 Db GCCCTCAGACAGGAGCTTACCTGCGCGCGGAGAACCTCTGGAATTCATCGAGATA 268

666 TTTGTCGTCGAGCGGCTCGCAAGTGTGAGAAAGCAATGAGTGAATTCATCGAGATA 725
 Db TTTGTCGTCGAGCGGCTCGCAAGTGTGAGAAAGCAATGAGTGAATTCATCGAGATA 328

726 GCGAGCTGACTGCAACAGTTCTCACCGTCATGCGCGCGGTTCTCCACCCCAAGTAAAT 785
 Db GCGAGCTGACTGCAACAGTTCTCACCGTCATGCGCGCGGTTCTCCACCCCAAGTAAAT 148

786 GATGGAATAGCTTGGCTGATGAATCAGAAGCGGCGAGCCCTCTCCGCGCATGAATCAGC 845
 Db GATGGAATAGCTTGGCTGATGAATCAGAAGCGGCGAGCCCTCTCCGCGCATGAATCAGC 88

846 GCGCTCGCTGTGAGCTCTGACCGGTAG 873
 Db GCGCTCGCTGTGAGCTCTGACCGGTAG 60

RESULT 6
 BD165655/c
 LOCUS BD165655
 DEFINITION Novel polynucleotide.
 ACCSSION BD165655
 VERSION BD165655.1 GI:27871467
 KEYWORDS JP 2002191370-A/3454.
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 627)

AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotide
 JOURNAL Patent: JP 2002191370-A 3454 09-JUL-2002;
 KYOMA HAKKO KOGYO CO LTD

COMMENT
 OS Corynebacterium glutamicum
 PN JP 2002191370-A/3454
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
 KEIKO OCHIAI,
 PI HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
 OZAKI
 PC C12N15/09, C12N15/00, C07K16/12, C07K16/40, C12M1/00, PC
 C12N1/15
 PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
 PC 04, C12P13/08,
 PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
 G01N33/566,
 PC G01N33/569, G01N33/69, G01N37/00//C12P21/08, (C12N1/21, C12R1:15),
 PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15),
 PC C12N15/00,
 PC C12N5/00, C12N15/00
 CC Novel polynucleotide
 FH Key Location/Qualifiers
 FT source 1..627
 /organism="Corynebacterium glutamicum".

FEATURES
 source 1..627
 Location/Qualifiers
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 65.1%; Score 568; DB 6; Length 627;
 Best Local Similarity 100.0%; Pred. No. 1.7e-116;
 Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

306 GTTTTCGCGGGTTTGAATCGGTGCGCTTCGCCCAAAATGTTGATGCGCGCGTGTGGGA 365
 Db GTTTTCGCGGGTTTGAATCGGTGCGCTTCGCCCAAAATGTTGATGCGCGCGTGTGGGA 568

366 AATCTCATCGATCGCTCCAACTCGCGGTGAGAAACTCCAAAGTTGTTGAGTGAATCAAG 425
 Db AATCTCATCGATCGCTCCAACTCGCGGTGAGAAACTCCAAAGTTGTTGAGTGAATCAAG 508

426 GCTGTTGTCAGCTGCTCAACTGACGAAGCACCATCAATGAGTCAATGTTGACGTTCAACAT 485
 Db GCTGTTGTCAGCTGCTCAACTGACGAAGCACCATCAATGAGTCAATGTTGACGTTCAACAT 448

486 GCGGTACTCTCTCTGTCGCGAGCACCATCGAAGCGCATCTGCGCAAGTGAATCAAT 545
 Db GCGGTACTCTCTCTGTCGCGAGCACCATCGAAGCGCATCTGCGCAAGTGAATCAAT 388

546 GCGTTCGCGGAGTCAATGAGCTGCGGAGCACCATCAATGAGTCAATGTTGACGTTCAACAT 605
 Db GCGTTCGCGGAGTCAATGAGCTGCGGAGCACCATCAATGAGTCAATGTTGACGTTCAACAT 328

606 GCCCTCAGACAGGAGCTTACCTGCGCGCGGAGAACCTCTGGAATTCATCGAGATA 665
 Db GCCCTCAGACAGGAGCTTACCTGCGCGCGGAGAACCTCTGGAATTCATCGAGATA 268

666 TTTGTCGTCGAGCGGCTCGCAAGTGTGAGAAAGCAATGAGTGAATTCATCGAGATA 725
 Db TTTGTCGTCGAGCGGCTCGCAAGTGTGAGAAAGCAATGAGTGAATTCATCGAGATA 208

726 GCGAGCTGACTGCAACAGTTCTCACCGTCATGCGCGCGGTTCTCCACCCCAAGTAAAT 785
 Db GCGAGCTGACTGCAACAGTTCTCACCGTCATGCGCGCGGTTCTCCACCCCAAGTAAAT 148

786 GATGGAATAGCTTGGCTGATGAATCAGAAGCGGCGAGCCCTCTCCGCGCATGAATCAGC 845
 Db GATGGAATAGCTTGGCTGATGAATCAGAAGCGGCGAGCCCTCTCCGCGCATGAATCAGC 88

Db 723 TTGTTCGGTGGACGGCCCTCGCGAAGTGGTGAGAAAGCAATGACGCCAAGACCAATTGTT 664
 QY 726 GCGAGCTGACTGCACAAAGATTCTCACCGTCAATCGCGCGGTTCCTCCACCCACGATTAAAT 785
 Db 663 GCGAGCTGACTGCACAAAGATTCTCACCGTCAATCGCGCGGTTCCTCCACCCACGATTAAAT 604
 QY 786 GATGCAATAGCTTGCTCATGATCAGAAAGCGGCGACGCCCTCTCCGGCCATGAACCTCAGC 845
 Db 603 GATGCAATAGCTTGCTCATGATCAGAAAGCGGCGACGCCCTCTCCGGCCATGAACCTCAGC 544
 QY 846 CGCCTCCGCTGTGAGCTCTGGACCGTAG 873
 Db 543 CGCCTCCGCTGTGAGCTCTGGACCGTAG 516

RESULT 8
 AX067087/c
 LOCUS
 DEFINITION
 TITLE
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AX067087 993 bp DNA linear PAT 24-JAN-2001
 Sequence 669 from Patent WO0100805.
 AX067087
 AX067087.1 GI:12544795

Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
 AUTHORS
 TITLE
 JOURNAL

Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.
 corynebacterium glutamicum genes encoding proteins involved in
 membrane synthesis and membrane transport
 Patent: WO 0100805-A 669 04-JAN-2001;
 BASF AKTIENGESSELLSCHAFT (DE)

FEATURES
 source

1..993
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101..973
 /note="unnamed protein product; RXN03164"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC26771.1"
 /db_xref="GI:12544796"
 /db_xref="REMBL:CAC26771"
 /translation="MIYRVGVSGILKPAISLGLMHNFGDDKPLSTORSIIHRAFDGSG
 VTHFDLANNYGPAPGASAEINFGRIREDLKSHEDELIISSKAGMDWMPGPGFGSGSG
 YLVSAGDQSLIRGLGLDVFDFHRRPDDPTLEETWALRIDIVASGKALYVIGISYSGP
 ELTAEEAEPMFEPCGLPIIHQSYSLINRWVEEPGCDLNLQANNGLGVAFSPFSE
 AQGLLDYKIDGIPGSGRASOGKSLSEGLMNVNINDVRKUNDIAOERGGSLAQMALA
 WVLREQREYGALEP"

CDS

ORIGIN

Query Match 46.3%; Score 404.2; DB 6; Length 993;
 Best local Similarity 98.8%; Pred. No. 76-80; 3; Indels 2; Gaps 1;
 Matches 418; Conservative 0; Mismatches

QY 453 AGCACAATCAATGCACTGGTCACGGTA--TCGCGCGCGTACTCTCTTGTCTGCGCAGC 510
 Db 993 AGCACCATCAATGCACTGGTCACGGTAATCGGCGCGGTACTCTCTTGTCTGCGCAGC 934
 QY 511 ACCCATGCAAGCGCCATCTGCGCAAGTGA CTGCCCGGCTTCTCGGCGATGTCAATTGAGC 570
 Db 933 ACCCATGCAAGCGCCATCTGCGCAAGTGA CTGCCCGGCTTCTCGGCGATGTCAATTGAGC 874
 QY 571 TTGCGGCACATATCAATATTGTCAGTTCAACTGCCCTCAGACAGAGCACTTACCTCGG 630
 Db 873 TTGCGGCACATATCAATATTGTCAGTTCAACTGCCCTCAGACAGAGCACTTACCTCGG 814
 QY 631 CTGGCGCGGAAACCTCTCGAAATTCACATCGAGATATTGTCCGTCGAGCAGGCCCTCGCGCA 690
 Db 813 CTGGCGCGGAAACCTCTCGAAATTCACATCGAGATATTGTCCGTCGAGCAGGCCCTCGCGCA 754

VHTDPELEQHGKRAVELADVELSEQTRTRWVSRVALIGRRPKFGRRTYIAPWHS
VHGITAGAGETNOTAELIAGFEDMRAADVAKELYGLPKPORDVAVAGELDDRLADIL
QELSDQALIELTIDIERAADVLEEMDDADLLGPELDPEDKAGVLLBELMDPEESAP
VRLMGFNPITGALMTDPLTDPSTTVAEALAMARNPDLPTSLASLVVVRPPPTAT
PTGKVLGVHLOKLLREPPSSLVGILLDPLDPLVANDSQETAARFATYNLVCGPVL
DENGHLGAVAVDDLDHMLPEDWADAGIRPGVKEETLG"
5264. 5782
/note="CE1185, similar to AE007003-6|AAK45530.1| percent
identity: 60 in 155 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17995.1"
/db_xref="GI:23493024"
/translation="MSMQATSGGNNRAELGLKPKPEGVGSPQTVBEAQAQVMDLS
DNEFFVAVLETIVGVLEVEVTGRLTWGRVLGAGASGAWLGLFGLVNGVLSNLF
SLLTGLVMGVVPGVLAAPVWASRGKRDFTSATQIVAGRYDILCRDPRAREARDMI
ARNMGLCHPRQ"
5305. 5500
/note="CE1186"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17996.1"
/db_xref="GI:23493025"
/translation="MRTTGYKALLIGLVAVLSSLLVAGSVDPNRPPEPRQVAEQK
ENLYATVDYLVGNVLSYLYTLVGNRGPVAVIRHEDNLIADFPDLNEGSDLVIG
CTGALLEVNVPLAGELEBYLAANVAGETDINSGRDRTFDALIGLSPRLDTADP
SNAQCPGANGLPQNVPIYRNPTFSRADREILNWSAGAITTSELQELBEAKELGRS
RELVEFVASKGG"
6660. 7373
/note="CE1187, similar to AE007262-3|AAK65408.1| percent
identity: 32 in 235 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17997.1"
/db_xref="GI:23493026"
/translation="METILIGVLLFVACVQALGFLGMLAAPILAMPDPLPCLLI
LLAFGLSLATWADRGVEVPMVWGLKRIKIPGVLGTVMAVMPVAGLKVLATAV
ILGTLSLVGSPGHRNSFLAGSGGFTTATGAPPLALNRSMSPERVGRGLS
VCFVGSISALITLGAAGLWMLQAAVLVAVPAVAVGLVGLSVGVNKLNRLLFLGS
VTLISLGSWITQAGIF"
7592. 8149
/note="CE1188"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17998.1"
/db_xref="GI:23493027"
/translation="MGRGDQFAERFPADPRCKRTIKWQKTKGALAAKAAALLLV
GGGTFMSNDEVTSQDAFAGNALSAADKWTNGQVVDNNAVAVKVPDGLV
FTQQISYETSGDNLSTLTTPGISTATADAALATLNDITITIDGLSELTSITG
SGNTYFEGVGRGVSALSIOQA"
complement (8190. 9443)
/note="CE1189, similar to AF024666-23|AAG03374.1| percent
identity: 84 in 408 aa"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC17999.1"
/db_xref="GI:23493028"
/translation="WAGPYSDPTTLDLLAQAQSPDLRMWLOGFINOILSTQADQ
VCGADYVQDRVNRNGRHRDFTVRGVIDVAVPKLTGSPFPDWLLERTRAE
ALTVTATCYLKGSTRMDLVATLGNNSKQSDMAKOLDQWEEFKRPLDTG
PLYVSDALTMKREGGVYKTSVLLATGVNAEGYRELLGMQVATSNVWGFGR
DLKAGNEVLVYDHALGIGTGAIVCEVLPNASQRCRTHFSKNLYGMVSKQWPTLS
AMPHETIQPDQAQVWGAQREVAFCQKFPDVADYLEALDELLAFNAPKSVTKV
WSNPTERANREIRRTVGIPEPDAVRLVGLVAEQHDDWIQQKRYSLTSLQ
TRAWTANIIDSDITSIIRKQVA"
complement (11439. 15143)
/gene="odha"
/protein_id="CAC25127.1"

Query Match 37.7%; Score 329.2; DB 1; Length 308650;
Best Local Similarity 74.4%; Pred. No. 3.2e-63;
Matches 415; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 316 GTTGTGAATTCGGTGGCTTCCGCCAAATTTGATGCGCGGCTGTGGGAAATCTCATCG 375
DB 187204 GTCTTGAATTCGGTGGCTTCCGCCAGATGTTTACCCCGCGTCAATGCGCATCTCATCG 187145
QY 376 ATCCGCTTCCAACTCCGCGCTCAGAAACTCAAGTTTGTGAGTGAATCAAGGCTGTCTCC 435
DB 187144 ATGCGCTTCCAGTTCCGCTCGGTGAATTCAGGTTGGCCAGGCGTCCGAGGTGGCATCG 187085
QY 436 AGCTGCTCAACTGACGAGACCAATCAATGCACTGGTCACTCCGCGCGCTTCTCT 495
DB 187084 AGCTGCTCACTGAGGACCACTGAGGCGGACCGTCCGAGGTGACCTCTCTCCACGAGGAG 187025
QY 496 CTTGTCTCGCGACGACCCATGCAAGCGCCATCTCGCAAGTGACTGCGCGGCTTCTCTGG 555
DB 187024 CCTGTCTCGCGACGACCCAGCGCATCTGCGCGACGACTGACCGCGCTCTCTGA 186965
QY 556 GCGATGCTTGAAGTTCGGGACCAATATCAATATGTTTCACTTCAACATGCGCTCAGAC 615
DB 186964 GCGATGCTTGAAGTTCGGGACGCGCCATCCGATGTTCTCTCGACGAGTGCCTTCGAC 186905
QY 616 AGGCACTTACCTGCTGCGCGGGAACCCCTCTGGAAATTCATCGAGATATTTTCGCTG 675
DB 186904 AGCGACTTGGCTTGGTGGCAGCGGACCCCTCGGCGACACCGTTCGAGTAGCGTTCGCTG 186845
QY 676 AGCAGGCGCTGCGCAAGTGGTGAGAAAGCAATGACGCAAGACCATTTGTTCGACTGAC 735
DB 186844 AGCAGGCGCTGCGCAAGTGGTGAGAAAGCAATGACGCAAGACCATTTGTTCGACTGAC 186785
QY 736 TCGACAGTTCCTACCGCTTCGCGGTTCTCTCACCAACGATTAAATGATGAATAG 795
DB 186784 TGAAGCAGGACTCCCTCTCATCCCGTTCTCTCACCGGTTCTCTCACCGGTTGACGAGGAGTAG 186725
QY 796 CTTGCTGATGAATCAGAAAGCGGCGAGCCCTCTCTCGGCATGAATCAAGCGCTTCGCT 855
DB 186724 CTCGCTGATGAATCAGAAAGCGGCGAGCCCTCTCTCGGCATGAATCAAGCGCTTCGCT 186665
QY 856 GTGAGCTCTGACCGTAG 873
DB 186664 GTGAGATCCGACCGTAG 186647
RESULT 10
AX063771 822 bp DNA linear PAT 24-JAN-2001
LOCUS Sequence 53 from Patent WO0100843.
DEFINITION AX063771
ACCESSION AX063771
VERSION AX063771.1 GI:12541483
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriidae; Corynebacterium.
REFERENCE 1
AUTHORS Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Habermeyer, G.
TITLE corynebacterium glutamicum genes encoding metabolic pathway
JOURNAL Patent: WO 0100843-A 53 04-JAN-2001;
BASF AKTIENGESSELLSCHAFT (DE)
FEATURES
source
1. 822
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101. 802
/note="unnamed protein product; RXA01394"
/codon_start=1
/transl_table=11
/protein_id="CAC25127.1"
CDS

AVVMTALA IKLMLMG"

.....

Db 406 GTCGATAGACGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 465
Qy 61 AACCCGAATGCGTATTGTCACGCGGTTTGTGTTTATCGGCGGCGTGGGCGCCAAATACGGC 120
Db 466 AACCCGAATGCGTATTGTCACGCGGTTTGTGTTTATCGGCGGCGTGGGCGCCAAATACGGC 525
Qy 121 GACACCGAGCGGTGATTTGCGCGCTGCGCGGTTTCGGCGAGCCTGATCTGCTTCCCG 180
Db 526 GACACCGAGCGGTGATTTGCGCGCTGCGCGGTTTCGGCGAGCCTGATCTGCTTCCCG 585
Qy 181 CTGCTGGGTTTCGGCGAGCAGCATTTGTCAGCCCGCTGTCAGCCCAAGCTGTGGCGC 240
Db 586 CTGCTGGGTTTCGGCGAGCAGCATTTGTCAGCCCGCTGTCAGCCCAAGCTGTGGCGC 645
Qy 241 TGGATCAACGTCGTGTCGCGAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
Db 646 TGGATCAACGTCGTGTCGCGAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 705
Qy 301 GGTAG 306
Db 706 GGTAG 711

RESULT 13
AX643030
LOCUS AX643030 712 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 9 from Patent EP1266986.
ACCESSION AX643030
VERSION AX643030.1 GI:28550160
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
AUTHORS Gunji, Y. and Yasueda, H.
TITLE Method for producing L-lysine or L-arginine by using methanol
JOURNAL assimilating bacterium
PATENT: EP 1266986-A 9 18-DEC-2002;
Ajinomoto Co., Inc. (JP)
FEATURES
source 1. .712
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
CDS 1. .375
/notes="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAD67812.1"
/db_xref="GI:28550161"
/translation="MWIMEIFITGLLGLSALLSIGPQNVIVKGIKRGELIANLVV
CLISDVFLFIATIGLVLLSNAPVLDIMRWGGIAYLLFWFAWAAKDMTNKVEAPQ
IIBETEPTVPDPLDPLGVFGGH"

ORIGIN
Query Match 34.9%; Score 304.4; DB 6; Length 712;
Best Local Similarity 99.7%; Pred. No. 1.5e-57;
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTCGATAGACGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 60
Db 407 GTCGATAGACGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 466
Qy 61 AACCCGAATGCGTATTGTCACGCGGTTTGTGTTTATCGGCGGCGTGGGCGCCAAATACGGC 120
Db 467 AACCCGAATGCGTATTGTCACGCGGTTTGTGTTTATCGGCGGCGTGGGCGCCAAATACGGC 526
Qy 121 GACACCGAGCGGTGATTTGCGCGCTGCGCGGTTTCGGCGAGCCTGATCTGCTTCCCG 180
Db 527 GACACCGAGCGGTGATTTGCGCGCTGCGCGGTTTCGGCGAGCCTGATCTGCTTCCCG 586

Qy 181 CTGCTGGGTTTCGGCGAGCAGCATTTGTCAGCCCGCTGTCAGCCCAAGCTGTGGCGC 240
Db 587 CTGCTGGGTTTCGGCGAGCAGCATTTGTCAGCCCGCTGTCAGCCCAAGCTGTGGCGC 646
Qy 241 TGGATCAACGTCGTGTCGCGAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
Db 647 TGGATCAACGTCGTGTCGCGAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 706
Qy 301 GGTAG 306
Db 707 GGTAG 712

RESULT 14
AX123539
LOCUS AX123539 708 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 3455 from Patent EP1108790.
ACCESSION AX123539
VERSION AX123539.1 GI:14041027
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 3455 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source 1. .708
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN
Query Match 34.7%; Score 303; DB 6; Length 708;
Best Local Similarity 100.0%; Pred. No. 3.1e-57;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGATAGACGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 60
Db 406 GTCGATAGACGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATGCTGCTGACCTGGTTG 465
Qy 61 AACCCGAATGCGTATTGTCAGCGGTTTGTGTTTATCGGCGGCGTGGGCGCCAAATACGGC 120
Db 466 AACCCGAATGCGTATTGTCAGCGGTTTGTGTTTATCGGCGGCGTGGGCGCCAAATACGGC 525
Qy 121 GACACCGAGCGGTGATTTGCGCGCTGCGCGGCTTTCGGCGAGCCTGATCTGCTTCCCG 180
Db 526 GACACCGAGCGGTGATTTGCGCGCTGCGCGGCTTTCGGCGAGCCTGATCTGCTTCCCG 585
Qy 181 CTGCTGGGTTTCGGCGAGCAGCATTTGTCAGCCCGCTGTCAGCCCAAGCTGTGGCGC 240
Db 586 CTGCTGGGTTTCGGCGAGCAGCATTTGTCAGCCCGCTGTCAGCCCAAGCTGTGGCGC 645
Qy 241 TGGATCAACGTCGTGTCGCGAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
Db 646 TGGATCAACGTCGTGTCGCGAGTTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 705
Qy 301 GGT 303
Db 706 GGT 708

RESULT 15
BD165656
LOCUS BD165656 708 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION BD165656
VERSION BD165656.1 GI:27871468
KEYWORDS JP 2002191370-A/3455.

SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 708)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayaishi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotide
JOURNAL Patent: JP 2002191370-A 3455 09-JUL-2002;
KIOWA HAKKO KOGYO CO LTD
COMMENT OS Corynebacterium glutamicum
PN JP 2002191370-A/3455
PD 09-JUL-2002
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HARUHIKO YOKOI,NAKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO
PI OZAKI
PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC
C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/3/
04,C12P13/08,
PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC
G01N33/566,
PC G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),
PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),
PC C12N15/00,
PC C12N5/00,C12N15/00
CC Novel polynucleotide
FH Key Location/Qualifiers
FT source 1..708
FT /organism='Corynebacterium glutamicum'.
FEATURES
source
1..708
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 34.7%; Score 303; DB 6; Length 708;
Best Local Similarity 100.0%; Pred. No. 3.1e-57;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGGTTG 60
DB 406 GTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGGTTG 465
QY 61 AACCCGATGCGTATTGGAGCGGTTTGTTGTTATCGCGCGGTGGCGCCCAATACGGC 120
DB 466 AACCCGATGCGTATTGGAGCGGTTTGTTGTTATCGCGCGGTGGCGCCCAATACGGC 525
QY 121 GACACCGAGCGGTGATTTTCGCGCTGGCGGTTTCGCGCAAGCTGATCTGGTTCGGC 180
DB 526 GACACCGAGCGGTGATTTTCGCGCTGGCGGTTTCGCGCAAGCTGATCTGGTTCGGC 585
QY 181 CTGGTGGTTTCGGCGCAGCAGCATTTGTCAGCCCGTGTTCAGCCCAAGGTGTGGCGC 240
DB 586 CTGGTGGTTTCGGCGCAGCAGCATTTGTCAGCCCGTGTTCAGCCCAAGGTGTGGCGC 645
QY 241 TGGATCAAGCTGCTGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATGTCGATG 300
DB 646 TGGATCAAGCTGCTGGCAGTGTGATGACCGCATTTGGCCATCAAACTGATGTCGATG 705
QY 301 GGT 303
DB 706 GGT 708

Search completed: March 15, 2004, 22:02:04
Job time : 3498.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 15, 2004, 10:09:44 ; Search time 355.983 Seconds
(without alignments)
10418.122 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1421_2293
Perfect score: 873
Sequence: 1 gtcgatagcagcgggttg.....ctgtgagctctggaccgtag 873

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2003as.*
8: geneseqn2003bs.*
9: geneseqn2003cs.*
10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	873	100.0	2374	2 AAT96816	Aat96816 DNA encod
2	873	100.0	2374	9 ADB66196	Adb66196 DNA fragm
C 3	873	100.0	349980	5 AAH68528	Aah68528 C Glutami
C 4	568	65.1	627	5 AAH68419	Aah68419 C Glutami
C 5	404.2	46.3	993	4 AAF68077	Aaf68077 Corynebac
C 6	326	37.3	822	4 AAF71779	Aaf71779 Corynebac
C 7	326	37.3	822	4 AAS96098	Aas96098 C. glutam
C 8	306	35.1	711	8 ACC80941	Acc80941 Lyse prot
C 9	304.4	34.9	712	8 ACC80942	Acc80942 Lyse24 pr
C 10	303	34.7	708	5 AAH68420	Aah68420 C Glutami
C 11	253.6	29.0	1095	7 ACA29651	Aca29651 Prokaryot
C 12	193.6	22.2	1568	4 AAH45375	Aah45375 C. thermo
C 13	105.8	12.1	1041	7 ACA25567	Aca25567 Prokaryot
C 14	96.8	11.1	1107	7 ACA26520	Aca26520 Prokaryot
C 15	96.2	11.0	1041	7 ACA24017	Aca24017 Prokaryot
C 16	80.4	9.2	1038	7 ACA32095	Aca32095 Prokaryot
C 17	79.4	9.1	1041	7 ACA19150	Aca19150 Prokaryot
C 18	78.2	9.0	1035	7 ACA45509	Aca45509 Prokaryot
C 19	76.4	8.8	999	7 ACA51519	Aca51519 Prokaryot
C 20	74.8	8.6	999	7 ACA50894	Aca50894 Prokaryot
C 21	71.6	8.2	5541	2 AAQ55755	Aaq55755 Escherich
C 22	69.2	7.9	349980	6 ABQ81844	Abq81844 Bifidobac
C 23	65.4	7.5	988	6 ABK72786	Abk72786 Bacillus

C 24	64.4	7.4	999	7 ACA24884	Aca24884 Prokaryot
C 25	59.2	6.8	990	7 ACA35500	Aca35500 Prokaryot
C 26	57.6	6.6	23128	4 AAS59552	Aas59552 Propionib
C 27	57.6	6.6	23128	7 ACF64481	Acf64481 Propionib
C 28	57	6.5	15239	2 AAT33536	Aat33536 BCG delet
C 29	57	6.5	110000	4 AA199682	Continuation (23 o
C 30	57	6.5	110000	4 AA199683	Continuation (23 o
C 31	56.4	6.5	636	3 AAS2691	Aas2691 Escherich
C 32	56.2	6.4	86114	6 ABX09143	Abx09143 Mycobacte
C 33	56.2	6.4	110000	4 AA199682	Continuation (6 of
C 34	56.2	6.4	110000	4 AA199683	Continuation (6 of
C 35	52.6	6.0	110000	6 ABA90521	ABA90521 20
C 36	52.4	6.0	990	7 ACA53775	ACA53775 Prokaryot
C 37	50.2	5.8	1002	9 ADE12817	Adel2817 DNA encod
C 38	50.2	5.8	1277	6 AAD31869	Aad31869 Lactobaci
C 39	48	5.5	2000	7 ADA71938	Ada71938 Rice gene
C 40	45.8	5.2	987	7 ACA33777	Aca33777 Prokaryot
C 41	42.4	4.9	2000	7 ADA71938	Ada71938 Rice gene
C 42	42.2	4.8	969	7 ACA26642	Aca26642 Prokaryot
C 43	42	4.8	349980	6 ABQ81847	Abq81847 Bifidobac
C 44	41	4.7	2157	4 AAH14096	Aah14096 Human cDN
C 45	41	4.7	3270	4 AAH14671	Aah14671 Human cDN

ALIGNMENTS

RESULT 1
AAT96816
ID AAT96816 standard; DNA; 2374 BP.
XX
AC AAT96816;
XX
DT 12-MAR-1998 (first entry)
XX
DE DNA encoding LysG, LysE and ORF3 from Corynebacterium glutamicum.
XX
KW LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;
Microbial production; amino acid; animal feed additive; ds.
XX
OS Corynebacterium glutamicum.
XX
FH Key Location/Qualifiers
FT CDS complement (82. .954)
FT /tag= a
FT /label= LysG
FT 1016. .1726
FT /tag= b
FT /label= LysE
FT complement (1723. .2373)
FT /tag= c
FT /label= orf3
DE19548222-AL.

26-JUN-1997.
22-DEC-1995; 95DE-01048222.
22-DEC-1995; 95DE-01048222.
(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
Vrijic M, Eggeling L, Sahn H;
WPI; 1997-333867/31.
P-PSDB; AAW37714, AAW37715, AAW37716.
Increasing microbial production of amino acids, especially lysine - by
PT improving export carrier activity or corresponding gene expression, also
PT new export and regulatory genes from Corynebacterium.
XX Claim 23 and 26; Page: 16pp; German.

XX CC This DNA, isolated from Corynebacterium glutamicum, contains the lysE, lysE and ORF3 genes. LysE and lysE encode a lysine transport regulatory protein and an export protein, respectively. Microbial production of amino acids (A) is improved by increasing the export-carrier activity and/or the export gene expression in a microorganism that produces (A). The method is specifically used to increase production of lysine, used as an animal feed additive. Other (A) are variously useful as pharmaceuticals, condiments and intermediates for fine chemicals. This method increases the amount of (A) secreted into the culture medium. Export of (A) has been found to depend on a single gene. NB. This sequence has been created from the information given in table 2 of the specification

XX SQ Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;

Query Match 100.0%; Score 873; DB 2; Length 2374; Best Local Similarity 100.0%; Pred. No. 4.9e-246; Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGATAGCAGCGGTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGCTGCTGGTTG 60
DB 1421 GTCGATAGCAGCGGTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGCTGCTGGTTG 1480

QY 61 AACCCGATGCGTATTGACGCGTTGTTGTTATCGCGCGCTGCGCGCAATACGCG 120
DB 1481 AACCCGATGCGTATTGACGCGTTGTTGTTATCGCGCGCTGCGCGCAATACGCG 1540

QY 121 GACCGGACGCGTGGATTTTCGCGCTGCGCGCTGCGCGCAATGCTGCTGCTGCTG 180
DB 1541 GACCGGACGCGTGGATTTTCGCGCTGCGCGCTGCGCGCAATGCTGCTGCTGCTG 1600

QY 181 CTGGTGGTTTCGCGCGAGCAGCATGTCAGCGCGCTGTCAGCGCGCAAGTGTGGCG 240
DB 1601 CTGGTGGTTTCGCGCGAGCAGCATGTCAGCGCGCTGTCAGCGCGCAAGTGTGGCG 1660

QY 241 TGGATCAAGTTCGCGCGAGCAGCATGTCAGCGCGCTGTCAGCGCGCAAGTGTGGCG 300
DB 1661 TGGATCAAGTTCGCGCGAGCAGCATGTCAGCGCGCTGTCAGCGCGCAAGTGTGGCG 1720

QY 301 GGTAGTTTCGCGCGTTTGGCAATCGGTGCGCTTCGCCCAATGTTGATGCGCGGTCG 360
DB 1721 GGTAGTTTCGCGCGTTTGGCAATCGGTGCGCTTCGCCCAATGTTGATGCGCGGTCG 1780

QY 361 TGGCAATTCATGATGCGCTTCAACTCGCGCTGAGCAAACTCCAGTTGTTGATGAA 420
DB 1781 TGGCAATTCATGATGCGCTTCAACTCGCGCTGAGCAAACTCCAGTTGTTGATGAA 1840

QY 421 TCAAGCTGTTGTCAGCTGCTCAACTGACGAGCAACCAATCAATGCACTGTCAGGTA 480
DB 1841 TCAAGCTGTTGTCAGCTGCTCAACTGACGAGCAACCAATCAATGCACTGTCAGGTA 1900

QY 481 TCGCGCGCTGATCTCTCTGTCGCGAGCAACCATGCAAGCGCATCTGCGCAATGAC 540
DB 1901 TCGCGCGCTGATCTCTCTGTCGCGAGCAACCATGCAAGCGCATCTGCGCAATGAC 1960

QY 541 TCGCGCGCTTCTGCGCGATGTCATGAGCTGCGGACCAATCAATGTTGTTGATG 600
DB 1961 TCGCGCGCTTCTGCGCGATGTCATGAGCTGCGGACCAATCAATGTTGTTGATG 2020

QY 601 AACATGCTCTGACAGGACTTACCTGCTGCGCGGAGCACTCTGGAATTCATGCG 660
DB 2021 AACATGCTCTGACAGGACTTACCTGCTGCGCGGAGCACTCTGGAATTCATGCG 2080

QY 661 AGATATTGTCGCGAGCGCTTCCGCAAGTGTGAGAAAGCAATGACGCCAAGCA 720
DB 2081 AGATATTGTCGCGAGCGCTTCCGCAAGTGTGAGAAAGCAATGACGCCAAGCA 2140

QY 721 TTGTTGCGAGCTGATGCAACAGTTCTCACCGTCATGCGCGGTTCTCCACCAAGCA 780
DB 2141 TTGTTGCGAGCTGATGCAACAGTTCTCACCGTCATGCGCGGTTCTCCACCAAGCA 2200

QY 781 TTAATGATGATAGTGGCTGATGAATCAGAAGCGGCGGCGCTCTCTCGCGCATGAAC 840

DB 2201 TTAATGATGATAGTGGCTGATGAATCAGAAGCGGCGGCGCTCTCTCGCGCATGAAC 2260

QY 841 TCAGCGCGCTTCCGCTGTGAGCTCTCGGACCGTAG 873

DB 2261 TCAGCGCGCTTCCGCTGTGAGCTCTCGGACCGTAG 2293

RESULT 2
ADB66196
ID ADB66196 standard; DNA; 2374 BP.
XX AC ADB66196;
XX DT 04-DEC-2003 (first entry)
XX DE DNA fragment containing C. glutamicum lysG and lysE genes.
XX KW L-arginine production; coryneform bacteria; lysE; arginine repressor;
XX KW argR; liver function promoting agent; amino acid infusion;
XX KW amino acid pharmaceutical; LysG; ds.
XX OS Corynebacterium glutamicum.
XX FH Key Location/Qualifiers
XX FT 1025..1726
XX FT /*tag= a
XX FT /product= "Protein encoded by lysE gene"
XX PN US2003113899-A1.
XX PD 19-JUN-2003.
XX PF 17-JUL-2002; 2002US-00196232.
XX PR 25-JUL-2001; 2001JP-00224586.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Yamaguchi M, Ito H, Gunji Y, Yasueda H;
XX DR WPI; 2003-708853/67.
XX DR P-PSDB; ADB66197.
XX PS A microorganism comprising enhanced expression of the lysE gene is useful for enhanced production of L-arginine.
XX Example 4; Page 26-27; 36pp; English.
XX The present invention relates to a method for producing L-arginine in a microorganism (e.g. coryneform bacteria) that has L-arginine producing ability and has been modified for enhanced expression of the lysE gene. The microorganism is also modified so that an arginine repressor (argR) does not function normally. The method of the invention is useful for the enhanced production of L-arginine which is useful in liver function promoting agents, amino acid infusion and comprehensive amino acid pharmaceuticals. The present sequence represents a DNA fragment containing Corynebacterium glutamicum LysG and LysE genes. Note: The present sequence is given as SEQ ID No:24 in the Sequence Listing but is referred to as SEQ ID No:25 in the rest of the specification.
XX SQ Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;

Query Match 100.0%; Score 873; DB 9; Length 2374; Best Local Similarity 100.0%; Pred. No. 4.9e-246; Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGATAGCAGCGGTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGCTGCTGGTTG 60
DB 1421 GTCGATAGCAGCGGTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGCTGCTGGTTG 1480

QY 61 AACCCGATGCGTATTGACGCGTTGTTGTTATCGCGCGCTGCGCGCAATACGCG 120

Db 1481 AACCCGATGCGTATTTGGACGGCTTTGTTTATCGCGCGCTCGCGCAATACGCG 1540
Qy 121 GACACCGGACGGTATTTGCGCGCTTGGCGGCTTCCGCGCAAGCTCTGATCTGGTTCCCG 180
Db 1541 GACACCGGACGGTATTTGCGCGCTTGGCGGCTTCCGCGCAAGCTCTGATCTGGTTCCCG 1600
Qy 181 CTGGTGGGTTTCCGCGCAGCAGCATTGTCACGCCGCTGTCACGCCCAAGGCTGTGCGCG 240
Db 1601 CTGGTGGGTTTCCGCGCAGCAGCATTGTCACGCCGCTGTCACGCCCAAGGCTGTGCGCG 1660
Qy 241 TGGATCAACGCTGCTGGCGCAGTGTGATGACCGCATGCGCCATCAAACTGATGTGATG 300
Db 1661 TGGATCAACGCTGCTGGCGCAGTGTGATGACCGCATGCGCCATCAAACTGATGTGATG 1720
Qy 301 GGTAGTGTTCGCGGTTTGGATCGGTGCGCTTCCGCCAATGTCATCGCGCGCTG 360
Db 1721 GGTAGTGTTCGCGGTTTGGATCGGTGCGCTTCCGCCAATGTCATCGCGCGCTG 1780
Qy 361 TGGGAATCTCATCGCTCGCTCAACTCGCGCTCAGAAATCTCAAGTTGTTGAGTGAA 420
Db 1781 TGGGAATCTCATCGCTCGCTCAACTCGCGCTCAGAAATCTCAAGTTGTTGAGTGAA 1840
Qy 421 TCAAGGCTGTTCAGCTGCTCAACTGACGAGCAATCAATGACATGCTGTCACGCTG 480
Db 1841 TCAAGGCTGTTCAGCTGCTCAACTGACGAGCAATCAATGACATGCTGTCACGCTG 1900
Qy 481 TCCGCGCGTACTCTCTCTGCTGCGCAGCAGCCATCAAGCGCATCTGCGCAAGTGAC 540
Db 1901 TCCGCGCGTACTCTCTCTGCTGCGCAGCAGCCATCAAGCGCATCTGCGCAAGTGAC 1960
Qy 541 TGCCCGGCTTCTCGCGCATGCTCAATGAGCTTGGGACCATATCAATATGTTTCAAGTTC 600
Db 1961 TGCCCGGCTTCTCGCGCATGCTCAATGAGCTTGGGACCATATCAATATGTTTCAAGTTC 2020
Qy 601 AACATGCTCTCAGACAGGACCTTACCTGCTGGCGCGGAAACCTCTGGAATTCATGCG 660
Db 2021 AACATGCTCTCAGACAGGACCTTACCTGCTGGCGCGGAAACCTCTGGAATTCATGCG 2080
Qy 661 AGATATTTGCTGCGCAGCAGCCCTGCGCAAGTGTGAGAAAGCAATGACGCAAGACCA 720
Db 2081 AGATATTTGCTGCGCAGCAGCCCTGCGCAAGTGTGAGAAAGCAATGACGCAAGACCA 2140
Qy 721 TTGTTGCGAGCTGACTGCAACCAAGTTCTCACCGTCTACCGCTTCTTCCACCCACCA 780
Db 2141 TTGTTGCGAGCTGACTGCAACCAAGTTCTCACCGTCTACCGCTTCTTCCACCCACCA 2200
Qy 781 TTAATGATGGAATGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 2201 TTAATGATGGAATGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2260
Qy 841 TCAGCGCCCTCGCGCTGCTGAGCTCTGAGACGCTAG 873
Db 2261 TCAGCGCCCTCGCGCTGCTGAGCTCTGAGACGCTAG 2293

RESULT 3
AAH68528/c
ID AAH68528 standard; DNA; 349980 BP.
XX AAH68528;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 7063.
XX
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EPI108790-A2.
XX
PD 20-JUN-2001.

XX
PP 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR MPI; 2001-376931/40.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Disclosure; SEQ ID NO 7063; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;
Query Match 100.0%; Score 873; DB 5; Length 349980;
Best Local Similarity 100.0%; Pred. No. 4e-245;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCCGATAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCGCAATCGTGTGACCTGGTTG 60
Db 128548 GTCCGATAGCAGCGGTTTGGGTAAAGCCCATGTTGATGCGCAATCGTGTGACCTGGTTG 128489
Qy 61 AACCCGATGCGTATTTGCGCGCGCTTTGTTTATCGCGCGCTGCGCGCAATACGCG 120
Db 128488 AACCCGATGCGTATTTGCGCGCGCTTTGTTTATCGCGCGCTGCGCGCAATACGCG 128429
Qy 121 GACACCGGACGGTATTTGCGCGCTTGGCGGCTTCCGCGCAAGCTCTGATCTGGTTCCCG 180
Db 128428 GACACCGGACGGTATTTGCGCGCTTGGCGGCTTCCGCGCAAGCTCTGATCTGGTTCCCG 128369
Qy 181 CTGGTGGGTTTCCGCGCAGCAGCATTGTCACGCCGCTGTCACGCCCAAGGCTGTGCGCG 240
Db 128368 CTGGTGGGTTTCCGCGCAGCAGCATTGTCACGCCGCTGTCACGCCCAAGGCTGTGCGCG 128309
Qy 241 TGGATCAACGCTGCTGGCGCAGTGTGATGACCGCATTGGCCATCAAACTGATGTGATG 300
Db 128308 TGGATCAACGCTGCTGGCGCAGTGTGATGACCGCATTGGCCATCAAACTGATGTGATG 128249
Qy 301 GGTAGTGTTCGCGGTTTGGATCGGTGCGCTTCCGCCAATGTCATCGCGCGCTG 360
Db 128248 GGTAGTGTTCGCGGTTTGGATCGGTGCGCTTCCGCCAATGTCATCGCGCGCTG 128189
Qy 361 TGGGAATCTCATCGCTCGCTCAACTCGCGCTCAGAAATCTCAAGTTGTTGAGTGAA 420
Db 128188 TGGGAATCTCATCGCTCGCTCAACTCGCGCTCAGAAATCTCAAGTTGTTGAGTGAA 128129
Qy 421 TCAAGGCTGTTCAGCTGCTCAACTGACGAGCAATCAATGACATGCTGTCACGCTG 480
Db 128128 TCAAGGCTGTTCAGCTGCTCAACTGACGAGCAATCAATGACATGCTGTCACGCTG 128069
Qy 481 TCAGCGCCCTCGCGCTGCTGAGCTCTGAGACGCTAG 540

Db 128068 TCCGCGCGTACTCTCTTCTGCTGCGCAGCACCCATGCAAGCGCATCTGCGCAAGTGAC 128009
QY 541 TCCCGCGCTTCTGCGCGATGTCATGAGCTTCCGACCAATATCAATATGTTTCAAGTTC 600
Db 128008 TGCCCGCGTCTCTGCGCGATGTCATGAGCTTCCGACCAATATCAATATGTTTCAAGTTC 127949
QY 601 AACATGCTCTCAGACAGGACATTTACCTGCTGCGCGGAGAACCTCTGGAATTCATCG 660
Db 127948 AACATGCTCTCAGACAGGACATTTACCTGCTGCGCGGAGAACCTCTGGAATTCATCG 127889
QY 661 AGATATTTGTCCTGAGAGCGCCCTGCGCAAGTGTGTGAGAAAGCAATGACGCAAGACCA 720
Db 127888 AGATATTTGTCCTGAGAGCGCCCTGCGCAAGTGTGTGAGAAAGCAATGACGCAAGACCA 127829
QY 721 TTGTTGGCAGCTGACTGCAACAAAGTTCTCACCGTATGCGCGGTCTCTCCACCAAGCA 780
Db 127828 TTGTTGGCAGCTGACTGCAACAAAGTTCTCACCGTATGCGCGGTCTCTCCACCAAGCA 127769
QY 781 TTAATGATGGAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCCGCGCATGAAC 840
Db 127768 TTAATGATGGAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCCGCGCATGAAC 127709
QY 841 TCAGCGCGCTCGCTGTGAGCTCTGACCGGTAG 873
Db 127708 TCAGCGCGCTCGCTGTGAGCTCTGACCGGTAG 127676

RESULT 4
AAH68419/c
ID AAH68419 standard; DNA; 627 BP.
XX
AC AAH68419;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 3454.
XX
DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW Coryneform bacterium; ds.
KW organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI: 2001-376931/40.
DR P-PSDB; AAG93200.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 1; SEQ ID NO 3454; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 627 BP; 139 A; 159 C; 185 G; 144 T; 0 U; 0 Other;
Query Match 65.1%; Score 568; DB 5; Length 627;
Best Local Similarity 100.0%; Pred. No. 1.5e-156;
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 306 GTTTTCGCGGGTTTGAATGCGTGGCTTCGCCAAATGTTGATGCGCGCTGCGGA 365
Db 627 GTTTTCGCGGGTTTGAATGCGTGGCTTCGCCAAATGTTGATGCGCGCTGCGGA 568
QY 366 AATCTCATGATGCGCTTCGAATGCGTGGCTTCGCCAAATGTTGATGCGGAATCAAG 425
Db 567 AATCTCATGATGCGCTTCGAATGCGTGGCTTCGCCAAATGTTGATGCGGAATCAAG 508
QY 426 GCTGTTGTTCCAGCTGCTCAACTGACGAAGCAACCAATCAATGCACTGCTGATCCGC 485
Db 507 GCTGTTGTTCCAGCTGCTCAACTGACGAAGCAACCAATCAATGCACTGCTGATCCGC 448
QY 486 GCGTATCTCTCTTCTGCTGCGGAGCAACCCATGCGAGGCGCATCTCGGCAAGTCACTGCC 545
Db 447 GCGTATCTCTCTTCTGCTGCGGAGCAACCCATGCGAGGCGCATCTCGGCAAGTCACTGCC 388
QY 546 GCGTCTCTGCGGATGTCATTGAGCTTTCGAGCAATATCAATATGTTTCAAGTTCACAT 605
Db 387 GCGTCTCTGCGGATGTCATTGAGCTTTCGAGCAATATCAATATGTTTCAAGTTCACAT 328
QY 606 GCGCTCAGACAGGACTTACCTTGGCTGCGGAGCAACCCCTCTGGAATTCATCGAGATA 665
Db 327 GCGCTCAGACAGGACTTACCTTGGCTGCGGAGCAACCCCTCTGGAATTCATCGAGATA 268
QY 666 TTTGTCGTCGAGAGCGCTTCGCGAGTGTGAGAAAGCAATGACGCAAGCATTTGT 725
Db 267 TTTGTCGTCGAGAGCGCTTCGCGAGTGTGAGAAAGCAATGACGCAAGCATTTGT 208
QY 726 GGCAGCTGACTGCAACAAAGTTCTCACCGTCTATCCCGGTTCTCCACCAAGATTAAT 785
Db 207 GGCAGCTGACTGCAACAAAGTTCTCACCGTCTATCCCGGTTCTCCACCAAGATTAAT 148
QY 786 GATGGAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCGGCATGAACTCAGC 845
Db 147 GATGGAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCGGCATGAACTCAGC 88
QY 846 GCGCTCGCTGTGAGCTCTGACCGGTAG 873
Db 87 GCGCTCGCTGTGAGCTCTGACCGGTAG 60

RESULT 5
AAH68077/c
ID AAH68077 standard; DNA; 993 BP.
XX
AC AAH68077;
XX
DT 11-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MCT protein encoding DNA SEQ ID NO:669.
XX
KW Corynebacterium glutamicum; brevivacterium lactofermentum; MCT;
KW membrane construction and membrane transport protein; petroleum spill;
KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
KW identification; microorganism; fine chemical production; transformation;
KW genome mapping; genetic engineering; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100805-A2.
XX

PD 04-JAN-2001.
 XX 23-JUN-2000; 2000WO-IB000926.
 PF 25-JUN-1999; 99US-0141031P.
 XX 08-JUL-1999; 99DE-01031454.
 PR 08-JUL-1999; 99DE-01031478.
 PR 08-JUL-1999; 99DE-01031563.
 PR 08-JUL-1999; 99DE-01032122.
 PR 09-JUL-1999; 99DE-01032124.
 PR 09-JUL-1999; 99DE-01032125.
 PR 09-JUL-1999; 99DE-01032128.
 PR 09-JUL-1999; 99DE-01032180.
 PR 09-JUL-1999; 99DE-01032182.
 PR 09-JUL-1999; 99DE-01032190.
 PR 09-JUL-1999; 99DE-01032191.
 PR 09-JUL-1999; 99DE-01032209.
 PR 09-JUL-1999; 99DE-01032212.
 PR 09-JUL-1999; 99DE-01032227.
 PR 09-JUL-1999; 99DE-01032228.
 PR 09-JUL-1999; 99DE-01032229.
 PR 09-JUL-1999; 99DE-01032230.
 PR 14-JUL-1999; 99DE-01032927.
 PR 14-JUL-1999; 99DE-01033005.
 PR 14-JUL-1999; 99DE-01033006.
 PR 27-AUG-1999; 99DE-01040764.
 PR 27-AUG-1999; 99DE-01040765.
 PR 27-AUG-1999; 99DE-01040766.
 PR 27-AUG-1999; 99DE-01040830.
 PR 27-AUG-1999; 99DE-01040831.
 PR 27-AUG-1999; 99DE-01040832.
 PR 27-AUG-1999; 99DE-01040833.
 PR 31-AUG-1999; 99DE-01041378.
 PR 31-AUG-1999; 99DE-01041379.
 PR 31-AUG-1999; 99DE-01041395.
 PR 03-SEP-1999; 99DE-01042077.
 PR 03-SEP-1999; 99DE-01042078.
 PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042088.
 XX (BADI) BASF AG.
 PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 XX WPI; 2001-071486/08.
 XX P-PSDB; AAB76844.
 PT Corynebacterium glutamicum nucleic acids encoding membrane construction
 PT and membrane transport proteins or their portions, useful for typing or
 PT identifying C. glutamicum or related bacteria, and as markers for
 PT transformation.
 XX Claim 3; Page 1107-1108; 1119pp; English.
 PS AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
 XX construction and membrane transport (MCT) proteins given in AAB76510 to
 CC AAB76847. The MCT nucleic acids and proteins are useful in the
 CC identification of microorganisms which can be used to produce fine
 CC chemicals, for modulating fine chemical production in C. glutamicum or
 CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
 CC identification of C. glutamicum or related bacteria, as reference points
 CC for mapping C. glutamicum genome, and as markers for transformation.
 CC AAF68082 and AAF68082 represent sequencing primers which are used in an
 CC example from the present invention
 XX SQ Sequence 993 BP; 222 A; 247 C; 277 G; 247 T; 0 U; 0 Other;
 Query Match 46.3%; Score 404.2; DB 4; Length 993;
 Best Local Similarity 98.8%; Pred. No. 2.7e-108;
 Matches 418; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
 QY 453 AGCACCAATCAATGCACTGGTACCGGTA--TCCGCGCCGTACTCTCTTGGCTCGCGCAGC 510
 ||||||||||||||||||||||||||| |||||||||||||||||||||||

Db 993 AGCACCAATCAATGCACTGGTACCGGTAATCCGGCGCGGTACTCTCTTTGCTCGCGCAGC 934
 QY 511 ACCCATGCAAGCCGATCTGGGCAAGTGAAGTCCGCGCGGTTCCTGGGCGATGTCAATGAGC 570
 Db 933 ACCCATGCAAGCCGATCTGGGCAAGTGAAGTCCGCGCGGTTCCTGGGCGATGTCAATGAGC 874
 QY 571 TTGCGGACCATATCAATATTGTTTCACTGTTCAACATGCGCCTCAGACAGGGAATACCTTGG 630
 Db 873 TTGCGGACCATATCAATATTGTTTCACTGTTCAACATGCGCCTCAGACAGGGAATACCTTGG 814
 QY 631 CTGGCGGGGAAACCTCTGGAATTCATTCAGATATTTGTCGAGAGGAGGCGCTTGGCGCA 690
 Db 813 CTGGCGGGGAAACCTCTGGAATTCATTCAGATATTTGTCGAGAGGAGGCGCTTGGCGCA 754
 QY 691 AGTGTGAGGAAGCAATGAGCGCAAGACCAATGTTGGCAGCTGACTGCAACAAAGTTCTCA 750
 Db 753 AGTGTGAGGAAGCAATGAGCGCAAGACCAATGTTGGCAGCTGACTGCAACAAAGTTCTCA 694
 QY 751 CCGTCATCGCCCGGTTCTCCACCAACGATTAATGATGAATAGCTTGGCTGATGAATC 810
 Db 693 CCGTCATCGCCCGGTTCTCCACCAACGATTAATGATGAATAGCTTGGCTGATGAATC 634
 QY 811 AGAAGCGGGCAGCGCTCTCCGCCATGAGTCTGAGCGGCTCGGCTGAGCTCTGAGCG 870
 Db 633 AGAAGCGGGCAGCGCTCTCCGCCATGAGTCTGAGCGGCTCGGCTGAGCTCTGAGCG 574
 QY 871 TAG 873
 Db 573 TAG 571
 RESULT 6
 AAF71779
 ID AAF71779 standard; DNA; 822 BP.
 XX
 AC AAF71779;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DE Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:53.
 XX
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
 KW fine chemical production; microorganism; organic acid; nucleoside;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN WO200100843-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-IB000923.
 XX
 PR 25-JUN-1999; 99US-0141031P.
 PR 01-JUL-1999; 99DE-01030476.
 PR 02-JUL-1999; 99US-0142101P.
 PR 08-JUL-1999; 99DE-01031415.
 PR 08-JUL-1999; 99DE-01031418.
 PR 08-JUL-1999; 99DE-01031419.
 PR 08-JUL-1999; 99DE-01031420.
 PR 08-JUL-1999; 99DE-01031424.
 PR 08-JUL-1999; 99DE-01031428.
 PR 08-JUL-1999; 99DE-01031434.
 PR 08-JUL-1999; 99DE-01031435.
 PR 08-JUL-1999; 99DE-01031443.
 PR 08-JUL-1999; 99DE-01031453.
 PR 08-JUL-1999; 99DE-01031457.
 PR 08-JUL-1999; 99DE-01031465.
 PR 08-JUL-1999; 99DE-01031478.
 PR 08-JUL-1999; 99DE-01031510.
 PR 08-JUL-1999; 99DE-01031541.

QY	61	AACCCGAATGCGTAATTTGGACGCGTTTGTGTTTATCGCGCGCGCTCGGCGGCAATACGGC	121
Db	557	AACCCGAATGCGTAATTTGGACGCGTTTGTGTTTATCGCGCGCGCTCGGCGGCAATACGGC	616
QY	121	GAACCGGACGCGTGGATATTTTCGCGCTGGCGGCTTCGGCGCAGCCTGATCTGGTCCCG	180
Db	617	GAACCGGACGCGTGGATATTTTCGCGCTGGCGGCTTCGGCGCAGCCTGATCTGGTCCCG	676
QY	181	CTGGTGGGTTTCGCGCGACGACGATCTGCACGCCCGCTGTCACAGCCCCCAAGGTGTGGCGC	240
Db	677	CTGGTGGGTTTCGCGCGACGACGATCTGTCACGCCCGCTGTCACAGCCCCCAAGGTGTGGCGC	736
QY	241	TGGATCAACGTCGTCGTGGCAGTTTGTGATGACGCAATTGGCCATCAAACTGATGTTGATG	300
Db	737	TGGATCAACGTCGTCGTGGCAGTTTGTGATGACGCAATTGGCCATCAAACTGATGTTGATG	796
QY	301	GGTTAGTTTCGCGCGGTTTTGGAATC	326
Db	797	GGTTAGTTTCGCGCGGTTTTGGAATC	822
RESULT 7			
AAS96098			
ID	AAS96098 standard; DNA; 822 BP.		
XX	AAS96098;		
DT	26-FEB-2002 (first entry)		
XX	C. glutamicum gene #23 encoding metabolic pathway protein.		
DE	Metabolic pathway protein; MP; lysine biosynthesis pathway;		
KW	methionine biosynthesis pathway; large-scale production of fine chemical;		
KW	Corynebacterium diphtheriae; diphtheria; ds.		
XX	Corynebacterium glutamicum.		
OS	Corynebacterium glutamicum.		
XX	WO200166573-A2.		
EN	13-SEP-2001.		
PD	22-DEC-2000; 2000WO-IB002035.		
XX	09-MAR-2000; 2000US-0187970P.		
PR	23-JUN-2000; 2000US-00606740.		
XX	(BADI) BASF AG.		
FA	Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;		
PI	Kim J, Lee H, Hwang B;		
XX	WPI; 2001-582269/65.		
DR	P-PSDB; AAU71888.		
XX	Nucleic acids encoding metabolic pathway proteins from Corynebacterium		
PT	glutamicum, useful for producing methionine and lysine in Corynebacterium		
PT	and Brevibacterium.		
XX	Disclosure; Page 214-215; 316pp; English.		
XX	The present invention relates to the isolation of novel Corynebacterium		
CC	glutamicum genes encoding metabolic pathway (MP) proteins (AAU71863-		
CC	AAU71922). The metabolic pathway proteins of the invention include		
CC	enzymes involved in the lysine and methionine biosynthetic pathways. The		
CC	polynucleotide sequences of the invention can be used for the large-scale		
CC	production and/or modulation of expression of fine chemicals such as		
CC	lysine and methionine. The sequences of the invention may be used to		
CC	identify C. glutamicum and related organisms e.g. C. diphtheriae in a		
CC	subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum		
CC	genes encoding the novel metabolic pathway proteins of the invention		
XX	Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;		
SQ			

Query Match 37.3%; Score 326; DB 4; Length 822;
 Best Local Similarity 100.0%; Pred. No. 2.5e-85; Mismatches 0; Indels 0; Gaps 0;
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGGTTG 60
 DB 497 GTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGGTTG 556
 QY 61 AACCCGATAGCGTATTCGACGCGTTTGTGTTATTCGCGCGCGTGGCGCCCAATAGCGC 120
 DB 557 AACCCGATAGCGTATTCGACGCGTTTGTGTTATTCGCGCGCGTGGCGCCCAATAGCGC 616
 QY 121 GACACCGAGCGGTGATTTTCGCGCTGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 180
 DB 617 GACACCGAGCGGTGATTTTCGCGCTGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 676
 QY 181 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCGTTCAGCGCCCAAGCGTGTGCGCG 240
 DB 677 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCGTTCAGCGCCCAAGCGTGTGCGCG 736
 QY 241 TGGATCAACGTCGTCTGCGCGAGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
 DB 737 TGGATCAACGTCGTCTGCGCGAGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 796
 QY 301 GGTAGTTTTCGCGGTTTTCGGAATC 326
 DB 797 GGTAGTTTTCGCGGTTTTCGGAATC 822

RESULT 8
 ACC80941
 ID ACC80941 standard; DNA; 711 BP.
 AC ACC80941;
 XX
 DT 27-OCT-2003 (revised)
 DT 11-AUG-2003 (first entry)
 XX
 DE LysE protein encoding sequence.
 XX
 KW L-lysine; L-arginine; LysE; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 FH Key Location/Qualifiers
 CDS 1..711
 FT /*tag= a
 FT /product= "lysE protein"
 XX
 PN EP1266966-A2.
 XX
 PD 18-DEC-2002.
 XX
 PF 05-JUN-2002; 2002EP-00012539.
 XX
 PR 12-JUN-2001; 2001JP-00177075.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 XX
 PI Gunji Y, Yasueda H;
 XX
 DR WPI; 2003-241171/24.
 DR P-PSDB; ABR58213.
 XX
 PT Novel DNA encoding variant of LysE protein from a coryneform bacterium,
 PT when introduced into methanol assimilating bacterium, facilitates
 PT excretion of L-lysine and/or L-arginine to outside of a cell.
 XX
 PS Example 1; Page 17-18; 23pp; English.
 XX
 CC The present invention relates to DNA encoding variants of protein with
 CC loop region and six hydrophobic helices which facilitates excretion of L-
 CC lysine and/or L-arginine to outside of cell of a methanol assimilating

CC bacterium when introduced into the bacterium. The method is used for
 CC encoding a protein which facilitates excretion of L-lysine, L-arginine or
 CC both of these L-amino acids to outside of a cell of a methanol
 CC assimilating bacterium when DNA of the method is introduced into the
 CC bacterium. The present sequence represents a lysE protein from
 CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003
 CC to standardise OS field)
 XX
 SQ Sequence 711 BP; 135 A; 173 C; 222 G; 181 T; 0 U; 0 Other;
 Query Match 35.1%; Score 306; DB 8; Length 711;
 Best Local Similarity 100.0%; Pred. No. 1.8e-79;
 Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGGTTG 60
 DB 406 CTCGATAAGCAGCGGTTTGGTAAAGCCCATGTTGATGCAATCGTCTGACCTGGTTG 465
 QY 61 AACCCGATAGCGTATTCGACGCGTTTGTGTTATTCGCGCGCGTGGCGCCCAATAGCGC 120
 DB 466 AACCCGATAGCGTATTCGACGCGTTTGTGTTATTCGCGCGCGTGGCGCCCAATAGCGC 525
 QY 121 GACACCGAGCGGTGATTTTCGCGCTGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 180
 DB 526 GACACCGAGCGGTGATTTTCGCGCTGCGCGTTCGCGCAAGCGTCTGCTGTTCCCG 585
 QY 181 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCGTTCAGCGCCCAAGCGTGTGCGCG 240
 DB 586 CTGCTGGGTTTCGCGCGCAGCAGCATTTGTCAAGCGCGTTCAGCGCCCAAGCGTGTGCGCG 645
 QY 241 TGGATCAACGTCGTCTGCGCGAGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
 DB 646 TGGATCAACGTCGTCTGCGCGAGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 705
 QY 301 GGTAG 306
 DB 706 GGTAG 711

RESULT 9
 ACC80942
 ID ACC80942 standard; DNA; 712 BP.
 AC ACC80942;
 XX
 DT 27-OCT-2003 (revised)
 DT 11-AUG-2003 (first entry)
 XX
 DE LysE24 protein encoding sequence.
 XX
 KW L-lysine; L-arginine; LysE24; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 FH Key Location/Qualifiers
 CDS 1..711
 FT /*tag= a
 FT /product= "lysE protein"
 XX
 PN EP1266966-A2.
 XX
 PD 18-DEC-2002.
 XX
 PF 05-JUN-2002; 2002EP-00012539.
 XX
 PR 12-JUN-2001; 2001JP-00177075.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 XX
 PI Gunji Y, Yasueda H;
 XX
 DR WPI; 2003-241171/24.
 DR P-PSDB; ABR58214.

XX Novel DNA encoding variant of LysE protein from a coryneform bacterium,
PT when introduced into methanol assimilating bacterium, facilitates
PT excretion of L-lysine and/or L-arginine to outside of a cell.
XX
XX Example 1; Page 19-20; 23pp; English.
PS
CC The present invention relates to DNA encoding variants of protein with
CC loop region and six hydrophobic helices which facilitates excretion of L-
CC lysine and/or L-arginine to outside of cell of a methanol assimilating
CC bacterium when introduced into the bacterium. The method is used for
CC encoding a protein which facilitates excretion of L-lysine, L-arginine or
CC both of these L-amino acids to outside of a cell of a methanol
CC assimilating bacterium when DNA of the method is introduced into the
CC bacterium. The present sequence represents a LysE24 protein from
CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003
CC to standardise OS field)
XX
SQ Sequence 712 BP; 133 A; 173 C; 224 G; 182 T; 0 U; 0 Other;

Query Match 34.9%; Score 304.4; DB 8; Length 712;
Best Local Similarity 99.7%; Pred. No. 5.3e-79;
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GTCCGATAGCAGCGGTTGGTAAAGCCATGCTGATGCGCAATCGTGTGACCTGGTTG 60
DB 407 GTCCGATAGCAGCGGTTGGTAAAGCCATGCTGATGCGCAATCGTGTGACCTGGTTG 466
QY 61 AACCCGATGCGTATTGGACGCGTTTGTGTTTATCGCGCGCGTTCGCGCAATACGCGC 120
DB 467 AACCCGATGCGTATTGGACGCGTTTGTGTTTATCGCGCGCGTTCGCGCAATACGCGC 526
QY 121 GACACCGAGCGGTGGATTTTCGCGCGCGGTTTCGCGCAATCGTGTGTTTCCG 180
DB 527 GACACCGAGCGGTGGATTTTCGCGCGCGGTTTCGCGCAATCGTGTGTTTCCG 586
QY 181 CTGCTGGGTTTCGCGCGCGAGCAGCATTTGTCACGCGCGGTTTCGCGCAATCGTGTGTTG 240
DB 587 CTGCTGGGTTTCGCGCGCGAGCAGCATTTGTCACGCGCGGTTTCGCGCAATCGTGTGTTG 646
QY 241 TGGATCAACGTCGTCGTCGCGAGTGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
DB 647 TGGATCAACGTCGTCGTCGCGAGTGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 706
QY 301 GGTAG 306
DB 707 GGTAG 712

RESULT 10
AAH68420
ID AAH68420 standard; DNA; 708 BP.
XX
AC AAH68420;
XX
DE 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 3455.
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX
XX Corynebacterium glutamicum.
OS
XX EP1108790-A2.
PN
XX 20-JUN-2001.
PD
XX 18-DEC-2000; 2000EP-00127688.
PF
XX 16-DEC-1999; 99JP-00377484.
PR
XX 07-APR-2000; 2000JP-00159162.
PR
XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.
PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai X, Yokoi H;
PI Tateishi N, Senoh A, Ikeda W, Ozaki A;
XX
XX WPI: 2001-376931/40.
DR P-PSDB; AAG93201.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.
PS
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids.
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;

Query Match 34.7%; Score 303; DB 5; Length 708;
Best Local Similarity 100.0%; Pred. No. 1.4e-78;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCCGATAGCAGCGGTTGGTAAAGCCATGCTGATGCGCAATCGTGTGACCTGGTTG 60
DB 406 GTCCGATAGCAGCGGTTGGTAAAGCCATGCTGATGCGCAATCGTGTGACCTGGTTG 465
QY 61 AACCCGATGCGTATTGGACGCGTTTGTGTTTATCGCGCGCGTTCGCGCAATACGCGC 120
DB 466 AACCCGATGCGTATTGGACGCGTTTGTGTTTATCGCGCGCGTTCGCGCAATACGCGC 525
QY 121 GACACCGAGCGGTGGATTTTCGCGCGCGGTTTCGCGCAATCGTGTGTTTCCG 180
DB 526 GACACCGAGCGGTGGATTTTCGCGCGCGGTTTCGCGCAATCGTGTGTTTCCG 585
QY 181 CTGCTGGGTTTCGCGCGCGAGCAGCATTTGTCACGCGCGGTTTCGCGCAATCGTGTGTTG 240
DB 586 CTGCTGGGTTTCGCGCGCGAGCAGCATTTGTCACGCGCGGTTTCGCGCAATCGTGTGTTG 645
QY 241 TGGATCAACGTCGTCGTCGCGAGTGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
DB 646 TGGATCAACGTCGTCGTCGCGAGTGTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 705
QY 301 GGT 303
DB 706 GGT 708
RESULT 11
ACA29651/c
ID ACA29651 standard; DNA; 1095 BP.
XX
XX ACA29651;
XX
XX 19-JUN-2003 (first entry)
DT
XX Prokaryotic essential gene #11308.
DE
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Corynebacterium diphtheriae.
OS

XX WO200277183-A2.
 XX
 XX
 XX 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 XX 25-OCT-2001; 2001US-0342923P.
 XX 08-FEB-2002; 2002US-00072851.
 XX 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX P-PSDB; ABU25781.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 XX for homologous nucleic acids required for cellular proliferation to
 XX isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 14; SEQ ID NO 17521; 1766pp; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 XX the 6213 antisense sequences given in the specification where expression
 XX of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX (1) a vector comprising a promoter operably linked to the nucleic acid
 XX encoding a polypeptide whose expression is inhibited by the antisense
 XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX polypeptide or its fragment whose expression is inhibited by the
 XX antisense nucleic acid; (4) an antibody capable of specifically binding
 XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX proliferation or the activity of a gene in an operon required for
 XX proliferation; (7) identifying a compound that influences the activity of
 XX the gene product or that has an activity against a biological pathway
 XX required for proliferation, or that inhibits cellular proliferation; (8)
 XX identifying a gene required for cellular proliferation or the biological
 XX pathway in which a proliferation-required gene or its gene product lies
 XX or a gene on which the test compound that inhibits proliferation of an
 XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX compound's activity; (11) a culture comprising strains in which the gene
 XX product is overexpressed or underexpressed; (12) determining the extent
 XX to which each of the strains is present in a culture or collection of
 XX strains; or (13) identifying the target of a compound that inhibits the
 XX proliferation of an organism. The antisense nucleic acids are useful for
 XX identifying proteins or screening for homologous nucleic acids required
 XX for cellular proliferation to isolate candidate molecules for rational
 XX drug discovery programs, or for screening homologous nucleic acids
 XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 XX prokaryotic essential genes. Note: The sequence data for this patent did
 XX not form part of the printed specification, but was obtained in
 XX electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1095 BP; 264 A; 247 C; 311 G; 273 T; 0 U; 0 Other;
 XX
 XX Query Match 29.0%; Score 253.6; DB 7; Length 1095;
 XX Best Local Similarity 65.6%; Pred. No. 5.5e-64;
 XX Matches 370; Conservative 0; Mismatches 194; Indels 0; Gaps 0;
 XX
 XX 310 TCGCGGTTTGAATCGTGGCTTCCGCAATGTTGAGCGCGGTCGGGAATC 369
 XX 1094 TCGTGAACCTTCGAGGAGTAGACACCGCCCAATATTAATTCAGCGCTTCGCAACA 1035
 XX 370 TCATCGATCGCTCCAACTCGCGGTCAGAAAACCTCAAGTTGTTGAGTCAATCAGGCTG 429
 XX 1034 GTGTCAATATAATGCGTTCTCTCAACGCTGAACGATTATTAAGAGCTCCAGATT 975

QY 430 TTGTCACAGTCTCACTGACGAAGACCAATCAATGCACTGGTCAAGGTATCCGCGCG 489
 DB 974 TGATCCAACTGCGTACCGAGATGCTCCGATCAAGCACTGGTACCGTTGTGCCCCA 915
 QY 490 TACTCTCTTGTCTGCGGAGCAACCCATGCAAGCGCATCTCGCAAGTGAAGTGGCCGGT 549
 DB 914 TAAATCGCTTGTCTCCGCGAGCACCCAGCAATAGCCAGCTGTGCCAACGTTTGGCCAGT 855
 QY 550 TCTTGGGCGATGTCATTGAGCTTGGGAGCCATATCAATATTTGTTCAAGTTCACATGCC 609
 DB 854 CTTAAGCGTAAATCATTCAGTGCAGCCACATGTCAAGATTTTGTAGGTTGAGCATCTCT 795
 QY 610 TCAGACAGGAGCTTACCTCTGCTGCGCGGGAACCTCTCGGAATTCATCGAGATTTTG 669
 DB 794 TTGCGCAGAGACTTCTCTGCTGCTGCGCGTGAATCACTGCGCAGCGCTCGAGATACCTG 735
 QY 670 TCGGTGAGCAGGCGCTTGGCGCAAGTGTGTGAGAAGCAATGACGCCAGACCATTTGTGCA 729
 DB 734 TCAGTTAAAGAGCCCTGTGCAAGTGGGGAATGCAATTAACCAACCCAGTCCGTTATTGCC 675
 QY 730 GCTGACTGCAACAAGTTCTCAACGCTCATGCGCGGTTCTCCACCCAAAGATTATGATG 789
 DB 674 GCGACTCTAACAGGTTCTCGCACTCTTCACTGCTTCTTACCCAGCGGTTCAAGATA 615
 QY 790 GAATAGCTTGGCTGATGAATCAGAAGCGGCGAGCCCTCTCCGCCATGAATCACTCAGCGCC 849
 DB 614 GAATAGCTCGGTTGATGAATCAGCAGCGCGGAGCCCTCAACGAGCTAGAAATTTCCGAGATC 555
 QY 850 TCGGCTGTGAGCTCTGAGCCGTAG 873
 DB 554 TCTCGGTTAGTTTCAGGTTCCATAG 531
 RESULT 12
 AH45375
 ID AH45375 standard; DNA; 1568 BP.
 AC AH45375;
 DT 11-SEP-2001 (first entry)
 DE C. thermoamino genes lysin biosynthetic enzyme lyse DNA.
 KW Heat-resistant; lysin biosynthesis; enzyme; coryneform;
 KW aspartate-semialdehyde dehydrogenase; lyse; ds.
 XX Corynebacterium thermoamino genes.
 XX JP2001120270-A.
 XX 08-MAY-2001.
 XX 01-NOV-1999; 99JP-003111148.
 XX 01-NOV-1999; 99JP-003111148.
 XX (AJIN) AJINOMOTO KK.
 XX WPI; 2001-364760/38.
 XX P-PSDB; AB64047.
 XX
 XX A heat-resistant lysin biosynthetic system enzyme gene of a high
 XX temperature-resistant coryneform microbe.
 XX Example 5; Page 22-24; 27pp; Japanese.
 XX The invention relates to a gene from a high temperature-resistant
 XX coryneform microbe that encodes a heat-resistant lysin biosynthetic
 XX enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and
 XX can be used for growing amino acid-producing microbes. The present
 XX sequence encodes an enzyme of the invention
 XX Sequence 1568 BP; 258 A; 525 C; 490 G; 295 T; 0 U; 0 Other;
 XX

	Query Match	22.2%;	Score 193.6;	DB 4;	Length 1568;
	Best Local Similarity	78.4%;	Pred. No. 2.9e-46;		
	Matches 232;	Conservative 0;	Mismatches 64;	Indels 0;	Gaps 0;
QY	10	CAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTCTGACCTGGTGTGAACCCGAAAT	69		
Db	1177	CGGCAGGTCCTGGGTCAGGCCCATGCTCATGAGCCATTGTCTGCACTGGCTCAATCCCCAAT	1236		
QY	70	GGGTATTGGACGCGTTTGTTGTTATCGCGCGCGCTCGGCGCGCAATACGGCGACACCCGGA	129		
Db	1237	GCCTACCTGGAATGCCCTTGGTCTTCATCGGTGGTGTGGAGCCCAAGTACGGGAGACCGGT	1296		
QY	130	CGGTGGATTTTCCGCGCTGGCGGCTTTCGGCGGAAGCCTGATCTGTTTCCCGCTGTGTGGGT	189		
Db	1297	CGGTGSAATCTTTCCTGCGGGTGCCCTTCGCGCGCAAGCCTGGTCTGGTTCCTCCACTGGTCTGGT	1356		
QY	190	TTTCGGCGGACAGCATTTGTACGCGCGCTGTTCAGCGCCCAAGGTGTGGCGCTCGATCAAC	249		
Db	1357	TACGGCGGGCGGCACTGTGCGCGTCCCTGTGCCAGCGCGGGCTCTGGCGGTGATCAAC	1416		
QY	250	GTGCTGCTGGCAGTTGTGATACCGCAATTTGGCCATCAAACTGATCTTGATGGGTGA	305		
Db	1417	ATAGCTGTGTGACCTGTGTGCTACCGGAATTTGGCGGTGAAGCTGATCTCTGATGGGTGA	1472		

RESULT 13
ACA25567/c
ID ACA25567 standard; DNA; 1041 BP.

XX	
DT	19-JUN-2003 (first entry)
XX	
DE	Prokaryotic essential gene #7224.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.

XX
OS Burkholderia fungorum.

XX
PN
WO200277183-A2.

XX
PD 03-007-2002-

XX
PF 21-MAR-2002: 2002WO-US009107.

XX
PR 21-MAR-2001: 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.

XX
C
(E) T E T A V A L M T N C

[illegible]

PI Wall D, Trawick JD, Carr

DR WPI; 2003-029926/02.

XX

for homologous nu

XX

CIAIII 14; SEX ID NO 131301 17006F; BIRTH-
FS
XX

CC the invention relates to an isolated nucleic acid where expression
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

polypeptide or its fragment whose expression is inhibited by the antiseptic nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antiseptic nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening candidate nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pat_sequences

Query Match	12.1%;	Score 105.8;	DB 7;	Length 1041;
Best Local Similarity	55.6%;	Pred. No. 1.6e-20;		
Matches 293;	Conservative 0;	Mismatches 192;	Indels 42;	Gaps 3

Qy	331	GCCTTCGCCCAATATGTTGATGCGCGCTGTGGGAAATCTCATGATGCGCTCCAACTCG	390
Dd	1022	GGCTTTTCCCAAGGTTGATGCGCCTTTCGGTGGGTAGCGGTGATCTCGCGCAGCTCG	963
Qy	391	GCCTCAGAAAACTCCAAGTTGTTGAGTGAATCAAGGCTGTGTTCAGAGTGTCTCAACTGAC	450
Dd	962	TCCCTTCGAGAACGCAAGGTTCTTCAGCGCGCGAAGTTTTTCAGGCACCTGTCTCGCGCG	903
Qy	451	GAAGCACCAATCAATGCACTGGTCAACGGTATCGCGCGGTACTCTCTTGTCTCGCGCAGC	510
Dd	902	CTCGCGCGATTCAGCACCGGAATCAGC-----CGCGATTCGGCGAGC	861
Qy	511	ACCCATGCAAGCGCATCTGCGCAAGTGACTGCGCGCGTTCCTGGCGGATGTCAATTGAGC	570
Dd	860	GCCACGGCGAGCGCCATTTGTGCGAGGCTCTGACGCGACGTTCGCGGATGTGTTGCTGAGC	801
Qy	571	TTGCGGACCATATCAATATATTGTTCACGTTCAACATGSCCTTCAGACAGGAGCTTACC----	626
Dd	800	TTGCGCACGTCTCGATATTCTTGGGGGCTCAGATGCTCTGCTTCAACGAACCGCGCCC	741
Qy	627	--CTGGCTGGCGGGAGAACCTCTTGGAATTCATCGAGATATTGTTCGTGAGCAGAGCCCC	684
Dd	740	GGCTTTTGACGCGCGGTTCTTCGCGCACCGGTTTCAGATACTGTCTGTGAGCAGAGCCCC	681
Qy	695	TGCGCAAGTGTGTGAGAAAGCAATAGCCCAAGACCAATTGTTGGCAGCTCACTGCAACAAG	744
Dd	690	TGAGCCAGCGCGGTGAACCGGATGCGCGCGCGCGGCTTTCTCCAGTGCGCCCGACAA--	622
Qy	745	TTCTCACCGTCATCGCCCGGTTCTCTCCACCCAAAGTAATGATGGAAATAGTTGGGCTGA	804
Dd	621	-----CTCATGCTCGATCCAGCGATTGAGCATGTTGTACGAAGGCTGA	579
Qy	805	TGAATCAGAGCGGCGAGCCCTCTTCGCCCATGACTCAGCGGCTTC	851
Dd	578	TGAATCAGCAGCGGACCTTTGATTTCGCGAGCAGAGCTTCGCCAATTC	532

RESULT 14
ACA26520/c
ID ACA26520 standard; DNA; 1107 BP.
XX

CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1041 BP; 193 A; 338 C; 350 G; 160 T; 0 U; 0 Other;

Query Match 11.0%; Score 96.2; DB 7; Length 1041;
 Best Local Similarity 55.7%; Pred. NO. 1.1e-17;
 Matches 241; Conservative 0; Mismatches 168; Indels 24; Gaps 2;
 QY 327 GGTGGCTTCGCGCAATGTTGATCGCGCGTGTGGAAATCTCATCGATCGCTCCAA 386
 DB 1026 GAGCGCTTTTCCACAGATTGATGCCGCTTTCGTCGCTAACGATCGATCTCCGAG 967
 QY 387 CTCGGGCTCAGAAATCCCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGCTCAAC 446
 DB 966 TTCTCCGCGAGAAATTCAGGTTCTGCAACGACGAGCTTCTCGCGACCTGCTCGGC 907
 QY 447 TGACGAAGCACCAATCAATCACTGCTACGCTATCGCGCGCTACTCTCTGCTCGCG 506
 DB 906 CCGGCTCGCGCGATCAGCGCGAGGTCAACGCGCGG-----TTACG 865
 QY 507 CAGCAGCATGCAAGCGCCATCTGCGAGTGAATGCTCGCGCGTCTCTGGCGATGTCATT 566
 DB 864 CAGCAGCATGCAAGCGCCATCTGCGAGTGAATGCTCGCGCGTCTCTGGCGATGTCATT 805
 QY 567 GAGCTTCGCGACCATATCAATTTGTTACGTTCAATGCGCTCAGACAGGAGCTTACC 626
 DB 804 GAGCTTCGCGACCATATCAATTTGTTACGTTCAATGCGCTCAGACAGGAGCTTACC 745
 QY 627 -----CTGGCTGGCGCGGAGCCCTCTGGATTCATCGAGATTTTTCCTGAGCAG 680
 DB 744 GCCCGGCTTGTTCACGCGCGCTCGCGCGGACGCGCTTCAGGTACTTCGACGAGAGCA 685
 QY 681 GCCCTGGCAGTGTGAGAAAGCAATGACGCAAGACCATTTGTCGAGCTGACTGCA 740
 DB 684 GCCCTGGCAGTGTGAGAAAGCAATGACGCAAGACCATTTGTCGAGCTGACTGCA 740
 QY 741 CAAGTCTCTACCG 753
 DB 624 CAGATCTCTTCG 612

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 15, 2004, 13:47:35 ; Search time 65.7645 Seconds
(without alignments)
7366.769 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1421_2293
Perfect score: 873
Sequence: 1 gtgcgaagcagcggttg.....ctgtgagctctggaccgtag 873

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCRU COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71.6	8.2	5541	1 US-08-920-812-20	Sequence 20, Appl
2	71.6	8.2	5541	1 US-08-920-827-20	Sequence 20, Appl
3	71.6	8.2	5541	1 US-08-921-177-20	Sequence 20, Appl
4	71.6	8.2	5541	1 US-08-362-577C-20	Sequence 20, Appl
5	71.6	8.2	5541	2 US-08-920-828-20	Sequence 20, Appl
6	61.8	7.1	699	4 US-09-252-991A-7908	Sequence 7908, Ap
7	61.8	7.1	834	4 US-09-252-991A-7643	Sequence 7643, Ap
8	61.8	7.1	894	4 US-09-252-991A-7841	Sequence 7841, Ap
9	61.2	7.0	1095	4 US-09-489-039A-5370	Sequence 5370, Ap
10	57	6.5	597	4 US-09-894-844-11	Sequence 11, Appl
11	57	6.5	15239	1 US-08-390-878-17	Sequence 17, Appl
12	57	6.5	4403765	3 US-09-103-840A-2	Sequence 2, Appl
13	57	6.5	4411529	3 US-09-103-840A-1	Sequence 1, Appl
14	53.8	6.2	1176	4 US-09-252-991A-7571	Sequence 7571, Ap
15	50.2	5.8	1002	4 US-09-724-623-24	Sequence 24, Appl
16	45	5.2	642	4 US-09-489-039A-4674	Sequence 4674, Ap
17	39.6	4.5	1935	2 US-08-492-027A-9	Sequence 9, Appl
18	37.6	4.3	4403765	3 US-09-103-840A-2	Sequence 2, Appl
19	37.6	4.3	4411529	3 US-09-103-840A-1	Sequence 1, Appl
20	36.2	4.1	1233	4 US-09-252-991A-2121	Sequence 2121, Ap
21	36.2	4.1	1395	4 US-09-252-991A-1954	Sequence 1954, Ap
22	36.2	4.1	1578	4 US-09-489-039A-3846	Sequence 3846, Ap
23	36.2	4.1	4056	4 US-09-489-039A-3861	Sequence 3861, Ap
24	35.4	4.1	1077	4 US-09-252-991A-14964	Sequence 14964, A
25	35.4	4.1	1155	4 US-09-252-991A-14787	Sequence 14787, A
26	35	4.0	1002	4 US-09-252-991A-1099	Sequence 1099, Ap
27	35	4.0	1092	4 US-09-252-991A-946	Sequence 946, Appl

C 28	35	4.0	1521	4 US-09-252-991A-989	Sequence 989, App
C 29	34.8	4.0	2730	4 US-09-489-039A-6744	Sequence 6744, Ap
C 30	34.6	4.0	417	4 US-09-252-991A-1351	Sequence 1351, Ap
C 31	34.6	4.0	1029	4 US-09-252-991A-1752	Sequence 1752, Ap
C 32	34.6	4.0	1098	4 US-09-252-991A-1304	Sequence 1304, Ap
C 33	34.6	4.0	1215	4 US-09-252-991A-1975	Sequence 1975, Ap
C 34	34.6	4.0	1230	4 US-09-252-991A-2136	Sequence 2136, Ap
C 35	34.6	4.0	1632	4 US-09-252-991A-1216	Sequence 1216, Ap
C 36	34.6	4.0	1746	4 US-09-252-991A-1163	Sequence 1163, Ap
C 37	34.6	4.0	1995	4 US-09-252-991A-14545	Sequence 14545, A
C 38	34.6	4.0	3396	4 US-09-252-991A-14676	Sequence 14676, A
C 39	34.6	4.0	3444	4 US-09-252-991A-15078	Sequence 15078, A
C 40	34.4	3.9	513	4 US-09-252-991A-2092	Sequence 2092, Ap
C 41	34.4	3.9	1500	4 US-09-252-991A-1868	Sequence 1868, Ap
C 42	34.4	3.9	1737	4 US-09-252-991A-1717	Sequence 1717, Ap
C 43	34.4	3.9	1767	4 US-09-252-991A-2007	Sequence 2007, Ap
C 44	34.4	3.9	2048	3 US-08-776-246-1	Sequence 1, Appli
C 45	34.4	3.9	2048	3 US-08-776-251-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-920-812-20
; Sequence 20, Application US/08920812
; Patent No. 5763188
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matshuisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,812
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: Clinical Isolate BC-625
; US-08-920-812-20

Query Match 8.2%; Score 71.6; DB 1; Length 5541;
Best Local Similarity 49.4%; Pred. No. 2.8e-11;
Matches 337; Conservative 0; Mismatches 299; Indels 46; Gaps 4;

203 CATTTGTCACGCGCTGTCAGCCCAAGGTGTGGCTGGATCAACGTCGTGGTGGCAG 262
Db CATATCATCTGACGCTTGTGCTCTCTGATTTGTATGGCTATCTTGACCGACTGAC 2746

263 TTGTGATGACCGATTTGGCCATCAAACTGATTTGATGGTGTAGTTTCGCGGGTTTGG 322
Db TGTCTATAATCACTGATCTTATGCGGCGCGGTTCTCGCGCCGTTTAAACAGGTCAAT 2806

323 ATCGGTGGCTTGGCCCAATGTTGATGCGCGCTGTGGGAAATCTCATCGCTGCT 382
Db ATTCGTAAGTGTAGGCTGTTGAGGTTGCGTTCGTTAGCAATTTTCGGCGTCAGACAC 3087

2807 TATCGGAGACGCTGCGCAAGATTGATGCGCTGCGGATATGCTGATCAATCTGG 2866

383 CCAACTGCGGCTGAGAACTCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGGCTGCT 442
Db CAGGCTCTCGGTGCTAAATGTCAGATTTTTCAGCGCTGCAAGTCTCTCAAGTTGTC 2926

443 CAACTGACGAGACCAATCAATGATGCTGATGCTGATGCTGATGCTGCTGCTGCT 502
Db CGCGCGGCTGCGCAACCAATCAATGATGCTGATGCTGATGCTGATGCTGATGCTGCT 2967

503 CGCGGAGACGCTGCGCAAGGCTGCTGCGGAGTCACTGCGCGCTGCTGCGGAGT 562
Db CTTTCAGCAACGCTTAAAGCAATTTGCGGCAATTTGCTGCGGCTGCTGCGGCTGCT 3027

563 CATTGAGCTTGGGACCATATCAATTTGTTACGTTTCAAGTTCAAATGCTGCGGAGT 622
Db CATTCAATAAGTGTAGGCTGTTGAGGTTGCGTTCGTTAGCAATTTTCGGCGTCAGACAC 3087

3028 CATTCAATAAGTGTAGGCTGTTGAGGTTGCGTTCGTTAGCAATTTTCGGCGTCAGACAC 3087

623 TACCTGGCTGGC-----GCGGGAACCTCTGGAATTCATTCAGATATTGCT 670
Db GAACTTTATTCCTTACGATGATCTCGGTAATCTTGGGAATGCGGTTGAGATATTTC 3147

671 CCGTGAGCGGCTTGGCGAAGTGTGGGAAAGCAATGACGCAAGACCATTTGTCGAG 730
Db CGGTGAGCAATCCCTGAGCGGAGGAGTAAAGCAATACAGCCCGCTTATTGTCAG 3207

3253 TGTACGAGGTTGATGAATTAACAGCGGAATTTTCCATCTCGCGGAGCAATCAACCAAT 3312

851 CGGCTGTGAGCTCTGGACGTA 872
Db TTTGGCTGCGCTCTGGCGAGTA 3334

RESULT 2
US-08-920-827-20
; Sequence 20, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Teuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,827
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 5541 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORGANISM: Escherichia coli
STRAIN: Clinical Isolate EC-625
US-08-920-827-20

Query Match 8.2%; Score 71.6; DB 1; Length 5541;
Best Local Similarity 49.4%; Pred. No. 2.8e-11;
Matches 337; Conservative 0; Mismatches 299; Indels 46; Gaps 4;

203 CATTGTGACGCGCTGTCAGCCCAAGGTGTGGCTGGATCAACGTCGTGGTGGCAG 262
Db CATTATCCATCTGACGCTTGTGCTCTCTGCAATTTGATGGGCTATCTTGACGACTGAC 2746

263 TTGTGATGACCGCATTTGGCCATCAAACTGATTTGATGGTGTAGTTTCGCGGGTTTGG 322
Db TGTCTATAATCACTGATCTTATGCGGCGCGGTTCTCGCGCCGTTTAAACAGGTCAAT 2806

323 AATCGGTGGCTTGGCCCAATGTTGATGCGCGCTGTGGGAAATCTCATCGCTGCT 382
Db TATCGGAGACGCTGCGCAAGATTGATGCGCTGCGGATATGCTGATCAATCTGG 2866

383 CCAACTGCGGCTGAGAACTCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGGCTGCT 442
Db CCAGCTCTCTCGGTGCTAAATGTCAGATTTTTCAGCGCTGCGAGCTTCTCTCTTGGT 2926

443 CAACTGACGAGCACCAATCAATGCACTGCTGACGCTATCCGCGCTGCTGCTTCTTGGT 502
Db CGCGCGGCTGCGCAACCAATCAATGCACTGCTGACGCTATCCGCGCTGCTGCTTCTTGGT 2967

503 CGCGGAGACGCTGCGCAAGGCTGCTGCGGAGTCACTGCGCGCTGCTGCGGAGT 562
Db CTTTCAGCAACGCTTAAAGCAATTTGCGGCAATTTGCTGCGGCTGCTGCGGCTGCT 3027

563 CATTGAGCTTGGGACCATATCAATTTGTTACGTTTCAAGTTCAAATGCTGCGGAGT 622
Db CATTCAATAAGTGTAGGCTGTTGAGGTTGCGTTCGTTAGCAATTTTCGGCGTCAGACAC 3087

3028 CATTCAATAAGTGTAGGCTGTTGAGGTTGCGTTCGTTAGCAATTTTCGGCGTCAGACAC 3087

623 TACCTGGCTGGC-----GCGGGAACCTCTGGAATTCATTCAGATATTGCT 670
Db GAACTTTATTCCTTACGATGATCTCGGTAATCTTGGGAATGCGGTTGAGATATTTC 3147

671 CCGTGAGCGGCTTGGCGAAGTGTGGGAAAGCAATGACGCAAGACCATTTGTCGAG 730
Db CGGTGAGCAATCCCTGAGCGGAGGAGTAAAGCAATACAGCCCGCTTATTGTCAG 3207

3148 CGGTGAGCAATCCCTGAGCGGAGGAGTAAAGCAATACAGCCCGCTTATTGTCAG 3207

731 CTGACTGCAACAGTTTTCACCGTCAATGCGCGGTTCTTCCACCAACGATTAATGATGG 790
Db TGTACGAGGTTGATGAATTAACAGCGGAATTTTCCATCTCGCGGAGCAATCAACCAAT 3312

851 CGGCTGTGAGCTCTGGACGTA 872
Db TTTGGCTGCGCTCTGGCGAGTA 3334

QY 791 AATAGCTTGGCTGATGATCAGAGCGGCGAGCCCTCTCCGCGCATGAACTCAGCGCGCT 850
Db 3253 TGTACGAGGTTGATGAATTAACAGCGGAATTTTCCACTCGCGCAGCAACTCAACCAATTT 3312
QY 851 CCGCTGTGAGCTCTCGACCGTA 872
Db 3313 TTTGCGTCCGCTCTGGCGAGTA 3334

RESULT 3
US-08-921-177-20
; Sequence 20, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORGANISM: Escherichia coli
; STAIN: Clinical isolate EC-625
US-08-921-177-20

Query Match 8.2%; Score 71.6; DB 1; Length 5541;
Best Local Similarity 49.4%; Pred. No. 2.8e-11;
Matches 337; Conservative 0; Mismatches 299; Indels 46; Gaps 4;

QY 203 CATTTGATCGCGCTGTCAGCGCCGAGGTGGCGTGGATCAACGTCGTCGTCGAG 262
Db 2687 CATTATCCATCTGACGCTTTGTGCTCTCTGCAATTTGATGGGCTATCTTGACCGAGTGAC 2746
QY 263 TTGTGATGACCGATTGGCCATCAAACTGATGTTGATGGGTAGTTTTCGCGGTTTGG 322
Db 2747 TCGTCATATCATGATCTTATGCGGCGCGGTTCTCGCGCCGCTTATTAACAGTCATT 2806
QY 323 AATCGGTGGCTTCGCCCAAAATGTTGATCGCGCGCTGCTGGGAAATCTCATGATCGCCT 382

Db 2807 TATCGAAGACGCGCTGCCACAGATTGATTCGCCATCGCGGATATGCTGATCAATCTGCG 2866
QY 383 CCAACTCGGCGTCAAGAAATCCCAAGTTTGTGAGTGAATCAAGCTGTTGTCAGGCTGCT 442
Db 2867 CAGCTCTCGTGGTGAATATCAGATTATTCAGGCGCTGACGTTCTCTCAAGTTGTC 2926
QY 443 CAACTGACGAAGCACCAATCAATGCTACTGTCACGGTATCGGCGCGTACTCTCTCTGCT 502
Db 2927 CGCGGGG-TGGCACTAATCAATACCGAGCTCACGGATCA-----T 2967
QY 503 CGCGGACACCCATGCAAGCGCCATCTGCGCAAGTGAAGTCTGCCCGGCTTCTCGGCGATGT 562
Db 2968 CTTTCAGCAACACAGCTTAAACGCCATTTGGGCCATTTGATTTGTCACGCTGCTGTCATT 3027
QY 563 CATTGAGCTTGGGACCATATCAATATTTGTTCACTGTTCAACATGCGCTCAGACGAGACT 622
Db 3028 CATTCAATAAGTGTAGGCTGTGAGGTTGGCTTCCGTAAGCATTTTGGCGTCAAGCAC 3087
QY 623 TACCTCGCTGGC-----GCGGAAACCTCTGGAATTCATCGAGATTTGT 670
Db 3088 GAACTTTATTCCTTCACGATGATCCGTGAATCTTGGGAATGCGGTTGAGATATTTTC 3147
QY 671 CCGTGAGAGCGCTCGCGCAAGTGTGAGAAAGCAATGACCGCAGACCACTTGTGGCAG 730
Db 3148 CCGTGAGCAATCTCTGAGCCAGAGAGTAAAGCAATACAGCCACGCGCTTATTTTGA 3207
QY 731 CTGACTGCAACAAAGTTCTCACCGCTATCGCGCGGTTCTTCCACCAACGATTAAATGATGG 790
Db 3208 GGGTATCAGCA-----GGCGGTTTATCCACCAGCGGTTCAATAAT 3252
QY 791 AATAGCTTGGCTGATGAATCAGAGCGGCGAGCGCTCTCGGCGATGAAGTCAAGCGCT 850
Db 3253 TGTACGAGGTTGATGAATTAACAGCGGAATTTTCCACTCGCGCAGCAACTCAACCAATT 3312
QY 851 CCGCTGTGAGCTCTGACCGTA 872
Db 3313 TTTGCGTCCGCTCTGGCGAGTA 3334

RESULT 4
US-08-362-577C-20
; Sequence 20, Application US/08362577C
; Patent No. 5807673
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,577C
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5541 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

PHAGE: T1
COLLECTOR TYPE: Genomic DNA

MOLECULE TYPE: GENOMIC DNA
ORIGINAL SOURCE:

ORIGINAL SOURCE: *Escherichia coli*

ORGANISM: C.
STRAIN: C.

Query Match 8.28: Score 71.6: DB 1: Length 5541;

Query Match

Best Local Similarity 43.4%, Recd. No. 2100
Matches 337: Conservative 0: Mismatches 299: Indels 46: Gaps 4:

203 CATTTGTCA CGCCCGCTGTCCA GCCCAAGTGTGGCGCTGGATCAACGTCGTCGTGGCAG 262

QY
203 CATGTGACGGCCCGCTGCAGCCCCCAAGGTGTGGCCTGGATGTCAGCCGCGCGGCGG

2697 C A T T T A T C C A T C T C A C C T T T C T C A T T T G T G A T G G G C T A T C T T G A C C G A C T G A C 2746

DB 2687 CATATCCATCTGACGTTGTGCTCTCTGCAATTTGTGATGGCTATCTTGACCGACCTGAC 2749

263 TTGTCGATGACCGCATTTGGCCATCAAACTGATGTTGATGGGTAGTTTTCGGGGGTTTGG 322

QY
263 TGTGATGACCGCATTTGGCCATCAATACAGTGTGATGCTTAGTATTCCGCCGTTTTCG

2747 ACCCTGCAAAATCTCATCTATATGCGCGCGCGCTTCTACGCGCGCTTATTACAGGTCAATT 2806

DB 2747 TCGTCATAATCACATGATCTTATGCGGGGCTCTCGGCCCGTAAATAACAGGICAT 2808

382

323 AATCGGTGGCCTTCGCCCAATGTTGATGCCGGCGTCTGGGAAATCTCATCGATCGCT 362

2956

Db 2807 TATCGGAAGACGCCCTGCCACAGATTTCAGCTCGCCATCGGCGATATGCTGATCAATCTGCG 2808

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

QY 383 CCAACTCGGCGTCAGAAAACCTCCAAGTTGTTGAGTGAATCAAGGCTGTGTCCAGCTGCT 442

2000

Db 2867 CCAGCTCCTCGGTGCTAAATGTCAGATTATTACAGCGCCTGCACGTTCTCCCTCAAGTTGTC 2926

[illegible]

QY 443 CAACTGACGAAGCACCAATCAATGCACCTGGTCACGGTATCCGGCCGTACTCTCCTTGCT 502

[illegible]

Db 2927 CCGCGGC-TGGACCAATCAATACCGAGTCACGGATCA-----T 2967

QY 503 CGCGAGCACCCATGCAAGCGCCATCTGCGCAAGTGACTGCCCGGTTCTGGCGGATGT 562

100

Db 2968 CTTTCAGCAACCAGCTTAACGCCATTGCGCCATTGATTTGTCCACGCTGCTGTGCCATT 3027

RESIST 5

RESULTS
IIS-08-920-828-20

US-08-520-828-20
. Sequence 20 Application IIS/08920828

; Sequence 20, Application
: Patient No 5953999

Patent No. 5853998

ADD TO NAME: Oboro Meadows

Db 3028 CATTCAATAGTGTAGCTGCTGAGCTTGGCTTCGGTAAGCAATTTTGGCGTCAGACCAC 3087
 QY 623 TACCTGGCTGGC-----GCGGAAACCTCTGGAATTCATCAGATATTTGT 670
 Db 3088 GAACCTTTATTCCTTCACGATGATCCGTGAATCTTCCGGAATGCCGTGAGATATTTTC 3147
 QY 671 CCGTGAGCAGGCTCGCGCAAGTGTGAGAAAGCAATGACGCAAGACATTTGTGGCAG 730
 Db 3148 CGTTCAGCAATCCCTGAGCAGGAGGTAAAGGCAATACAGCCAGCGCGTATTTTGA 3207
 QY 731 CTGATGCAACAAGTTCTCCCGTCAATCCCGGTTCTTCCACCAACGATTAATGATGG 790
 Db 3208 GGGTATCCAGCA-----GCGCGCTTTTATCCACCCAGCGGTTCAATAAT 3252
 QY 791 AATAGTCTGGCTGATGATCAGAAAGCGGCGAGCCCTCTCCGCAATGACTCAGCGCCT 850
 Db 3253 TGTACGAAGGTTTGAATTAACAGCGGAATTTTCCACTCGCGCAACTCAACCATTT 3312
 QY 851 CCGCTGTGAGCTCTGACCCGTA 872
 Db 3313 TTGCGTCTGCTCTGGCGAGTA 3334

RESULT 6
 US-09-252-991A-7908
 ; Sequence 7908, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 7908
 ; LENGTH: 699
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-7908

Query Match 7.1%; Score 61.8; DB 4; Length 699;
 Best Local Similarity 54.6%; Pred. No. 8.9e-09;
 Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;
 QY 35 TGATGGCAATCGTCTGACCTGTTGAACCCGAATGGGTATTTGGACGCGTTTGTGTTA 94
 Db 434 TGGCGGCACTGGCGGTACCTGCTCAACCCCGACGCTATCTCGATACCGTACTTCTGA 493
 QY 95 TCGCGGCGTTCGCGCGCAATACGCGGCAACCGGACGCGTGGATTTTCGCGCGTGGCGGT 154
 Db 494 TCGGTTCTGCTCGCGCGCCAGCAGG-----CGCGCGCGCGCGCGCTATGCGCTCGCGCGG 547
 QY 155 TCGCGGCAAGCTGATCTGCTTCCGCTGGTGGTTCGCGCGCAGCAGCATTTGTACGCC 214
 Db 548 CACGCGCTCTGCTGATGTTCTTCGCGCTCGCGCGCGCGGATGGCTGGCGCCCT 607
 QY 215 GCCTGTCCAGCCCAAGGTGGCGGTGATCAAGCTGCTGCGCGCAGCAGCATTTGTACCG 274
 Db 608 GGCTGGCGCGCGCGCGACCTGGCGCGCTGCTCGACCTGATGTTGGCGCGCATGATGCTGG 667
 QY 275 CATTGGCCATCAAACTGATGTTGATGGTTA 305
 Db 668 GCATGGCGCGGCAACTGCTGTTCCGGGATA 698

RESULT 7
 US-09-252-991A-7643/c

; Sequence 7643, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 7643
 ; LENGTH: 834
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-7643

Query Match 7.1%; Score 61.8; DB 4; Length 834;
 Best Local Similarity 54.6%; Pred. No. 9.7e-09;
 Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;
 QY 35 TGATGGCAATCGTCTGACCTGTTGAACCCGAATGGGTATTTGGACGCGTTTGTGTTA 94
 Db 386 TGGCGGCACTGGCGGTACCTGCTCAACCCCGACGCTATCTCGATACCGTACTTCTGA 327
 QY 95 TCGCGGCGTTCGCGCGCAATACGCGGCAACCGGACGCGTGGATTTTCGCGCGTGGCGGT 154
 Db 326 TCGGTTCTGCTCGCGCGCCAGCAGG-----CGCGCGCGCGCGCGCTATGCTTCGCGCGG 273
 QY 155 TCGCGGCAAGCTGATCTGCTTCCGCTGGTGGTTCGCGCGCAGCAGCATTTGTACGCC 214
 Db 272 CACGCGCTCTGATGTTCTTCGCGCTCGCGCGCGCGCATGCTGGCGCCCT 213
 QY 215 GCCTGTCCAGCCCAAGGTGGCGGTGATCAAGCTGCTGCGCGCAGTGTGATGACCG 274
 Db 212 GCTGGCGCGCGCGCGCACCTGCGCGCTGCTGACCTGATGTTGGCGCGCATGATGCTGG 153
 QY 275 CATTGGCCATCAAACTGATGTTGATGGTTA 305
 Db 152 GCATGGCGCGGCAACTGCTGTTCCGGGATA 122

RESULT 8
 US-09-252-991A-7841
 ; Sequence 7841, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 7841
 ; LENGTH: 894
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-7841

Query Match 7.1%; Score 61.8; DB 4; Length 894;
 Best Local Similarity 54.6%; Pred. No. 1e-08;
 Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;
 QY 35 TGATGGCAATCGTCTGACCTGTTGAACCCGAATGGGTATTTGGACGCGTTTGTGTTA 94

```

Db 339 TGGCGGCACTGGCGGTACCTGCTCAACCCCGACGCTATCTCGATACCGTACTTCTGA 398
QY 95 TCGCGCGCGTTCGCGCGCAAPACGGGCAACCGGAGGTGATTTTCGGCGCGTGGCGGT 154
Db 399 TCGGTTCGCTCGCGCGCCGACGAG-----CGCGCGCGCGCGCTATGCGCTTCGGCGCG 452
QY 155 TCGCGGCAAGCTGATCTGTTTCCGCTGCTGGTTCGGCGCGCGAGCATTTGTCAAGCC 214
Db 453 CACGCGCTCGCTGATGTTCTTTCGCTCGCCCTCGCGCGCGGATGGTTCGCGCCCT 512
QY 215 CGCTGTCCAGCCCCAAGGTGTGGCTGGATCAAGCTGCTGGCGAGTTGTGATGACGG 274
Db 513 GCTGGCGCGCGCGCCACCTGGCGCTGCTGACCTGATGTTGGCGCGCATGATGCTGG 572
QY 275 CATTGGCCATCAACTGATGTTGATGGTTA 305
Db 573 GCATGGCGCGCAACTGCTGTTCCGGGGATA 603

RESULT 9
US-09-489-039A-5370/c
; Sequence 5370, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5370
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5370

Query Match 7.0%; Score 61.2; DB 4; Length 1095;
Best Local Similarity 51.2%; Pred. No. 1.7e-08;
Matches 209; Conservative 0; Mismatches 178; Indels 21; Gaps 2;

QY 469 CTGGTCACGGTATCCGCGCGTACTCTCTCTGCTCGCGCAGCACCCATGCAAGCGCCATC 528
Db 998 CTGGCGCGATCAGCAGCGAGGTGACTTCTCTTCTGGCGCAACCCAGCGCCAGCGCCATC 939
QY 529 TCGGCAAGTGAATGCGCGGCTCTCGGCGATGTCATTGAGTTGGGACCATATCAATA 588
Db 938 TCGCACAGCTTCTGCGCCACGCGCTTCCGCTGCTTAAAGCTGACGAATTTTCTTAAT 879
QY 589 TTGTTTCACTTCAACATGCGCTCAGACAGGAGTACTTACCTGGCT---GGCGGGGAACCC 645
Db 878 CGGGTGGAGTCACTGCTCGCGGTGCGAAGAGAGTCTGCTGGCGCGCGCGAGTCTG 819
QY 646 TCTGGAATTCACATGAGATATTTGCTCGTGGAGAGCGCTCGCGAAGTGTGAGAAACA 705
Db 818 GCGGGAATGCGTTGAGATAGCGTCTGTAAGTTGCGCGCGCGCGCGCGTGAACGCT 759
QY 706 ATGACGCCAGACCATTTGTTGGAGCTGACCTGACCAAGTTCTTACGCTATCGCCCGGT 765
Db 758 ATGCTTCCAAATCCCTCCGCTCTGCAAAAATCCAGCAATCC-----T 717
QY 766 TCTTCCACCAACCATTAATGATGAATAGCTTGGCTGATGAATCAGAAAGCGGCGGCC 825
Db 716 TCTTCCACCGCGGTTTGAACATCGAGTAGGAGGCTGGTGAATGATGAGAGGGGTGCG 657
QY 826 TCTTCCGCGCATGAATCAGCGCTCCGCTGCTGAGCTCTGGACCGTAG 873
Db 656 AGGTGCTTAAGATCTTCAACGCGCTCCGCGCGCTGGCCAGAGGATAG 609

RESULT 10

```

```

US-09-894-844-11
; Sequence 11, Application US/09894844
; Patent No. 688166
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; TITLE OF INVENTION: the M. Tuberculosis Complex
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/09/894,844
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-09-894-844-11

Query Match 6.5%; Score 57; DB 4; Length 597;
Best Local Similarity 53.1%; Pred. No. 2.3e-07;
Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 26 AGCCATCTTGATGGCAATCGTGTGCTGACCTGTTGACCCGATGCTATTTGGACGCT 85
Db 326 AGGTCTCTGTGACCTGTGGCGCATTCACGTTCTCAACCCACACGCTCTACTCGACCG 385
QY 86 TTGTGTTATTCGGCGCGTTCGGCGCAATACCGCGCAATACCGCACACCGGATTTTCGCG 145
Db 386 TGTGTGTTAGCGCGCTGGCCACGAGCACAGCA---CCAGCGCTGGCTTTCGCGC 442
QY 146 CTGGCGGCTTCGGCGCAAGCCTGATCTGCTTCCGCTGGTGGTTCGGCGCAGCAGCAT 205
Db 443 TCGGGCGGCTCACAGCAGTGGCGGTATGTTTCGCCACCCCTCGGGTTCGGAGCGCGCG 502
QY 206 TGTACAGCGCCCTGTCACGCGCCCAAGGTGCGGCTGGATCAAGCTGCTGTCGAGTTG 265
Db 503 TGGCGGGCTGTTTCCACCAACCCCGCTGCTGGAGAAATCTCTGACGGCTGATCGGCTCA 562
QY 266 TGATGACCGCATTTGGCCATCAAACTGATGTTGA 298
Db 563 TGATGTTGGCTGGGAATCTGCTGACCGTGA 595

RESULT 11
US-08-390-878-17/c
; Sequence 17, Application US/08390878
; Patent No. 5700683
; GENERAL INFORMATION:
; APPLICANT: Stover, Charles K.
; APPLICANT: Mahairas, Gregory G.
; TITLE OF INVENTION: VIRULENCE-ATTENUATING GENETIC DELETIONS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower, 20th
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/390,878
 ; FILING DATE: 17-FEB-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hunter, Tom
 ; REGISTRATION NUMBER: 38,498
 ; REFERENCE/DOCKET NUMBER: 15371A-17
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/543/9600
 ; TELEFAX: 415/543/5043
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 15239 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-390-878-17

Query Match 6.5%; Score 57; DB 1; Length 15239;
 Best Local Similarity 53.1%; Pred. No. 1.1e-06;
 Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;
 QY 26 AGCCCATGTTGATGCAATGCTGACCTGTTGAACCGCAATGCGTATTGGACGCGT 85
 Db |||||
 QY 4792 AGTCTGTTGACCTGTGCGGCATTACGTTCTCAACCCACACGTTACTCGACACG 4733
 Db |||||
 QY 86 TTGTGTTTATCGCGCGCGTGGCGGCAATACGCGACACCGGAGCGTGGATTTCGCG 145
 Db |||||
 QY 4732 TCGTGTGTTAGGCGCGTGGCCACAGACACAGCA---CCAGCGCTGGCTGTTCGGCC 4676
 Db |||||
 QY 146 CTGGCGGTTGCGGCAAGCCTGATCTGCTTCCCGCTGGTGGGTTTCGGCGGACGACAT 205
 Db |||||
 QY 4675 TCGGCGGTTGCGGCAAGCCTGATCTGCTTCCCGCTGGTGGGTTTCGGCGGACGACAT 4616
 Db |||||
 QY 206 TGTCAAGCGCGTGTGTCAGACCCCAAGGTGTGCGCTGATCAACGTCGTCGTCGCGAGTTG 265
 Db |||||
 QY 4615 TCGCGCGGCTGTTCAACCAACCGCGCTGTGAGAAATCTCGACGCGCTGATCGCGTCA 4556
 Db |||||
 QY 266 TGATGACCGCATTTGCCATCAAACTGATGTGA 298
 Db |||||
 QY 4555 TGATGTTGCGTGGGAATCTCGTGACCGTGA 4523
 Db |||||

RESULT 12
 US-09-103-840A-2
 ; Sequence 2, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 4403765
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; FEATURE:
 ; OTHER INFORMATION: CDC 1551
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence
 ; OTHER INFORMATION: represent a, t, c or g
 ; US-09-103-840A-2

Query Match 6.5%; Score 57; DB 3; Length 4403765;
 Best Local Similarity 53.1%; Pred. No. 1.7e-05;
 Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 26 AGCCCATGTTGATGCAATGCTGACCTGTTGAACCGCAATGCGTATTGGACGCGT 85
 Db |||||
 QY 2227633 AGTCTGTTGACCTGTGCGGCATTACGTTCTCAACCCACACGTTACTCGACACG 2227692
 Db |||||
 QY 86 TTGTGTTTATCGCGCGCGTGGCGGCAATACGCGACACCGGAGCGTGGATTTCGCG 145
 Db |||||
 QY 2227693 TCGTGTGTTAGGCGCGCTGGCCAAAGACACAGCA---CCAGCGCTGGCTGTTCGGCC 2227749
 Db |||||
 QY 146 CTGGCGGTTGCGGCAAGCCTGATCTGCTTCCCGCTGGTGGGTTTCGGCGGACGACAT 205
 Db |||||
 QY 2227750 TCGGCGGCGGTACAGCGGATGTTGCGCACCCCTCGGTTTCGGAGCGCGCGGT 2227809
 Db |||||
 QY 206 TGTCAAGCGCGTGTGTCAGACCCCAAGGTGTGCGCTGATCAACGTCGTCGTCGCGAGTTG 265
 Db |||||
 QY 2227810 TCGCGCGGCTGTTTCAACCAACCGCGCTGTGAGAAATCTCGACGCGCTGATCGCGTCA 2227869
 Db |||||
 QY 266 TGATGACCGCATTTGCCATCAAACTGATGTGA 298
 Db |||||
 QY 2227870 TGATGTTGCGTGGGAATCTCGTGACCGTGA 2227902
 Db |||||

RESULT 13
 US-09-103-840A-1
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 ; US-09-103-840A-1

Query Match 6.5%; Score 57; DB 3; Length 4411529;
 Best Local Similarity 53.1%; Pred. No. 1.7e-05;
 Matches 145; Conservative 0; Mismatches 125; Indels 3; Gaps 1;
 QY 26 AGCCCATGTTGATGCAATGCTGACCTGTTGAACCGCAATGCGTATTGGACGCGT 85
 Db |||||
 QY 2230334 AGTCTGTTGACCTGTGCGGCATTACGTTCTCAACCCACACGTTACTCGACACG 2230393
 Db |||||
 QY 86 TTGTGTTTATCGCGCGCGTGGCGGCAATACGCGACACCGGAGCGTGGATTTCGCG 145
 Db |||||
 QY 2230394 TCGTGTGTTAGGCGCGTGGCCAAAGACACAGCA---CCAGCGCTGGCTGTTCGGCC 2230450
 Db |||||
 QY 146 CTGGCGGTTGCGGCAAGCCTGATCTGCTTCCCGCTGGTGGGTTTCGGCGGACGACAT 205
 Db |||||
 QY 2230451 TCGGCGGCGTACAGCCAGTGGGTATGTTTGGCCACCCCTCGGTTTCGGAGCGCGCGT 2230510
 Db |||||
 QY 206 TGTCAAGCGCGTGTGTCAGACCCCAAGGTGTGCGCTGATCAACGTCGTCGTCGCGAGTTG 265
 Db |||||
 QY 2230511 TCGCGGCGTGTTCACCAACCCCGCTGTGAGAAATCTCGTGACCGCTGATCGCGTCA 2230570
 Db |||||
 QY 266 TGATGACCGCATTTGCCATCAAACTGATGTGA 298
 Db |||||
 QY 2230571 TGATGTTGCGTGGGAATCTCGTGACCGTGA 2230603
 Db |||||

RESULT 14
 US-09-252-991A-7571/c
 ; Sequence 7571, Application US/09252991A

Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 7571
 LENGTH: 1176
 TYPE: DNA
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-7571

Query Match 6.2%; Score 53.8; DB 4; Length 1176;
 Best Local Similarity 54.3%; Pred. No. 2.9e-06;
 Matches 134; Conservative 0; Mismatches 107; Indels 6; Gaps 1;
 QY 59 TGAACCCGAATGCGTATTTGGACGCGTTTGTATTATCGGCGGCTCGCGCGCAATACG 118
 DB 1176 TCAACCCCGACGCTATCTCGATACCGTACTTCTGATCGGTTCTGCTCGGCGCCAGCAGG 1117
 QY 119 GCGACACCGGACGCGTGAATTTTCGCGCGTGGCGGTTTCGCGGCAAGCTGATCTGTTCC 178
 DB 1116 -----CGGCGCGGCGGCTATGCGCTTCGCGCGCGCCAGCGCTCGCTGATGTTCT 1063
 QY 179 CGCTGGTGGGTTTCGCGCGACGAGATTGTTCAGCGCGCTCTCCAGCGCCCAAGGTGTGGC 238
 DB 1062 TGCGCTCGCCCTCGGCGGCGCATGGCTGGCGCCCTTGGCTGGCGCGCCGCGCCACTGGC 1003
 QY 239 GCTGGATCAACGTCGTCGGGAGTGTGTGATGACCGCATGCGCCATCAAACTGATGTGA 298
 DB 1002 GCTGCTCGACCTGATGTTGGCGGCGCATGATGCTGGCGCATGCGCGCGCAACTGCTGTTCC 943
 QY 299 TGGGTTA 305
 DB 942 GGGGATA 936

RESULT 15
 US-09-724-623-24/c
 Sequence 24, Application US/09724623
 Patent No. 6476209
 GENERAL INFORMATION:
 APPLICANT: Glenn, Matthew
 APPLICANT: Lubbers, Mark W
 APPLICANT: Dekker, James
 TITLE OF INVENTION: Polynucleotides, materials incorporating
 TITLE OF INVENTION: them, and methods for using them.
 FILE REFERENCE: 1048U1
 CURRENT APPLICATION NUMBER: US/09/724,623
 CURRENT FILING DATE: 2000-11-28
 NUMBER OF SEQ ID NOS: 124
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 24
 LENGTH: 1002
 TYPE: DNA
 ORGANISM: Lactobacillus rhamnosus
 US-09-724-623-24

Query Match 5.8%; Score 50.2; DB 4; Length 1002;
 Best Local Similarity 49.3%; Pred. No. 3.2e-05;
 Matches 176; Conservative 0; Mismatches 163; Indels 18; Gaps 1;
 QY 365 AAATCTCATCGATCGCTCCAACTCGGCGTCAGAAACTCCAGTTGTTGAGTGAATCAA 424
 DB 994 AAATATCATCAATTTGTTGAATCTCTTCAGCAGTAAGGTGAGATGTTCCGTTGCTTTAA 935

QY 425 GGCTGTGTCGAGTGTCACTGACGAGCAACCAATCAATGCACTGCTCAGGTATCCG 484
 DB 934 GGTATCTGAGGTGTCACTGAGGTGTGCTCCCAATGATCAGCTTGTGACAA-----881
 QY 485 CGCGGTACTCTCTCTGCTCGCGACGACCCCAATGCAAGCGCCATCTGCGCAAGTACTGCC 544
 DB 880 -----CGGATCCCGTAAACCAACCAAGCCCAATTTGACTCAGGGTTTGGT 833
 QY 545 CGCGTTCCTGGCGGATGTCAATTGAGCTTTCGGGACCATATCAATATTCTTCAGTTCAACA 604
 DB 832 CAGGATCATGCGCAATTTCAATTAAGCGCATTTAGTTGCTTAAACCAAGCCCTTTTGGCCT 773
 QY 605 TGCCTCTCAGACAGGGACTTACCCCTGGCTGGCGGGAACCCCTCTGGAATTCATCGAGAT 664
 DB 772 TAGCAAAAGTGGCCTTTTGGTTGGATGATTTTGAAGTATCCGGAATTCCTTTAGGT 713
 QY 665 ATTTGTCGTGAGCAGGCCCTTCGCGAGTGTGTGAGAAAGCATGACGCCAAGCCAT 721
 DB 712 AGCGATCTGATTAACAAGCCCTTCTGATTAACGCTCCGTATGCAATCAACCCCGGACCAT 656

Search completed: March 16, 2004, 03:01:13
 Job time : 82.7645 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2004, 00:51:03 ; Search time 312.967 Seconds
(without alignments)
10269.549 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1421_2293

Perfect score: 873

Sequence: 1 gtcgataagcagcggttg.....ctgtgagctctggacgtag 873

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 184079884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	873	100.0	2374	14	US-10-196-232-24	Sequence 24, Appl
2	873	100.0	3309400	9	US-09-738-626-1	Sequence 1, Appl
3	568	65.1	697	9	US-09-738-626-3454	Sequence 3454, Ap
4	404.2	46.3	923	12	US-10-627-476-669	Sequence 669, App
5	326	37.3	822	10	US-09-746-660A-51	Sequence 51, Appl
6	306	35.1	711	14	US-10-166-142-7	Sequence 7, Appl
7	304.4	34.9	712	14	US-10-166-142-9	Sequence 9, Appl
8	303	34.7	708	9	US-09-738-626-3455	Sequence 3455, Ap
9	253.6	29.0	1095	12	US-10-282-122A-17521	Sequence 17521, A
10	136.6	15.6	1026	15	US-10-369-493-35374	Sequence 35374, A
11	136.6	15.6	1026	15	US-10-369-493-38826	Sequence 38826, A
12	136.6	15.6	1028	15	US-10-369-493-38124	Sequence 38124, A
13	105.8	12.1	1041	12	US-10-282-122A-13437	Sequence 13437, A
14	96.8	11.1	1107	12	US-10-282-122A-14390	Sequence 14390, A
15	96.2	11.0	1041	12	US-10-282-122A-11887	Sequence 11887, A

c 16	90.4	10.4	1026	15	US-10-369-493-37651	Sequence 37651, A
c 17	89.6	10.3	993	14	US-10-156-761-1151	Sequence 1151, Ap
c 18	89.6	10.3	9025608	14	US-10-156-761-1	Sequence 1, Appli
c 19	80.4	9.2	1038	12	US-10-282-122A-19965	Sequence 19965, A
c 20	79.4	9.1	1041	12	US-10-282-122A-7020	Sequence 7020, Ap
c 21	78.2	9.0	1035	12	US-10-282-122A-33379	Sequence 33379, A
c 22	76.4	8.8	999	12	US-10-282-122A-39389	Sequence 39389, A
c 23	74.8	8.6	999	12	US-10-282-122A-38764	Sequence 38764, A
c 24	70.8	8.1	745	15	US-10-260-238-1450	Sequence 1450, Ap
c 25	65.4	7.5	988	9	US-09-974-300-77	Sequence 77, Appl
c 26	64.4	7.4	999	12	US-10-282-122A-12754	Sequence 12754, A
c 27	62.8	7.2	609	14	US-10-156-761-1109	Sequence 1109, Ap
c 28	62.8	7.2	9025608	14	US-10-156-761-1	Sequence 1, Appli
c 29	59.2	6.8	990	12	US-10-282-122A-23370	Sequence 23370, A
c 30	57	6.5	597	9	US-09-894-844-11	Sequence 11, Appl
c 31	57	6.5	597	9	US-10-388-902-11	Sequence 11, Appl
c 32	56.2	6.4	86114	14	US-10-080-170-648	Sequence 648, App
c 33	52.4	6.0	990	12	US-10-282-122A-41645	Sequence 41645, A
c 34	50.2	5.8	1002	12	US-10-282-122A-930-24	Sequence 24, Appl
c 35	50.2	5.8	1277	9	US-09-971-536-15	Sequence 15, Appl
c 36	47	5.4	927	15	US-10-369-493-28453	Sequence 28453, A
c 37	47	5.4	927	15	US-10-369-493-31213	Sequence 31213, A
c 38	45.8	5.2	987	12	US-10-282-122A-21647	Sequence 21647, A
c 39	42.2	4.8	969	12	US-10-282-122A-14512	Sequence 14512, A
c 40	41	4.7	8473	16	US-10-038-854-39	Sequence 39, Appl
c 41	41	4.7	8487	16	US-10-038-854-41	Sequence 41, Appl
c 42	41	4.7	8645	16	US-10-038-854-37	Sequence 37, Appl
c 43	41	4.7	8675	16	US-10-038-854-35	Sequence 35, Appl
c 44	40.4	4.6	1392	15	US-10-369-493-31368	Sequence 31368, A
c 45	40.4	4.6	1443	15	US-10-369-493-28610	Sequence 28610, A

ALIGNMENTS

RESULT 1
US-10-196-232-24
; Sequence 24, Application US/10196232
; Publication No. US20030113899A1
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, MIKIKO
; APPLICANT: ITO, HISAO
; APPLICANT: GUNJI, YOSHIYA
; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: 225391USO
; CURRENT APPLICATION NUMBER: US/10/196,232
; CURRENT FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: JP 2001-224586
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 2374
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1025)..(1723)
; OTHER INFORMATION:
US-10-196-232-24

Query Match 100.0%; Score 873; DB 14; Length 2374;
Best Local Similarity 100.0%; Pred. No. 1.2e-272;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GTCCATAGCAGCGGTTTGGTAAAGCCCATGTTGATGCGCAATCGTGTGACCTGTTG	60
Db	1421	GTCCATAGCAGCGGTTTGGTAAAGCCCATGTTGATGCGCAATCGTGTGACCTGTTG	1480
Qy	61	AACCCGAAATCGTATTTCGACGCGTGTGTTTATCGGCGCGCTCGGCGCGCAATACGGC	120
Db	1481	AACCCGAAATCGTATTTCGACGCGTGTGTTTATCGGCGCGCTCGGCGCGCAATACGGC	1540

QY	121	GACCCGCGACGGTGGATTTTCGCGCTGGCGGTTTCGCGCAAGCCTGATCTGGTTCGCG	180
Db	1541	GACCCGCGACGGTGGATTTTCGCGCTGGCGGTTTCGCGCAAGCCTGATCTGGTTCGCG	1600
QY	181	CTGGTGGGTTTCGCGCGCAGCAGCATGTTCAAGCCCGCTGTCCAGCCCCAAGGTGTGGCGC	240
Db	1601	CTGGTGGGTTTCGCGCGCAGCAGCATGTTCAAGCCCGCTGTCCAGCCCCAAGGTGTGGCGC	1660
QY	241	TGGATCAACGTCGTGCTGGCAGTTGTGATGACCGGCATGGCCATCMAACTGATGTGATG	300
Db	1661	TGGATCAACGTCGTGCTGGCAGTTGTGATGACCGGCATGGCCATCMAACTGATGTGATG	1720
QY	301	GGTTAGTTTTCGCGGGGTTTGGAAATGGTGGGCTTCGCCCAAACTGTGATGCCGCGCTCG	360
Db	1721	GGTTAGTTTTCGCGGGGTTTGGAAATGGTGGGCTTCGCCCAAACTGTGATGCCGCGCTCG	1780
QY	361	TGGAAATCTCATCGATCGCCTCCAACTCGGCGTCAGAAACTCCAAAGTTTGTGAGTGAA	420
Db	1781	TGGAAATCTCATCGATCGCCTCCAACTCGGCGTCAGAAACTCCAAAGTTTGTGAGTGAA	1840
QY	421	TCAAGGCTGTGTCTCAGCTGTCTCAACTGACGAGCACAATCAATGACTGTGTCAGGTA	480
Db	1841	TCAAGGCTGTGTCTCAGCTGTCTCAACTGACGAGCACAATCAATGACTGTGTCAGGTA	1900
QY	481	TCCGCGCGCTACTCTCTTGTCTCGCGCAGCAGCCATCGAGCGGCATCTCGCGCAAGTCAC	540
Db	1901	TCCGCGCGCTACTCTCTTGTCTCGCGCAGCAGCCATCGAGCGGCATCTCGCGCAAGTCAC	1960
QY	541	TGCGCGCTTCTGGCGGATGTCAATTGAGCTTGGCGACCATCAATAATTTGTTACGTTTC	600
Db	1961	TGCGCGCTTCTGGCGGATGTCAATTGAGCTTGGCGACCATCAATAATTTGTTACGTTTC	2020
QY	601	AACATGCCCTCAGACAGGCACTTACCTGCTGGCGCGGGAACCTCTGGAAATCCATCG	660
Db	2021	AACATGCCCTCAGACAGGCACTTACCTGCTGGCGCGGGAACCTCTGGAAATCCATCG	2080
QY	661	AGATATTTGTCGTGACGAGGCCCTTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCA	720
Db	2081	AGATATTTGTCGTGACGAGGCCCTTGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCA	2140
QY	721	TTGTTGGCAGCTGACTGCAACAGTTCTCACCGTCAATCGCCCGTTCTCTCACCCAAAGA	780
Db	2141	TTGTTGGCAGCTGACTGCAACAGTTCTCACCGTCAATCGCCCGTTCTCTCACCCAAAGA	2200
QY	781	TTAATGATGGAATAGCTTGGCTGTGATGAATCAGAAGCGGCGAGCCCTCTCTCGGCATGAAC	840
Db	2201	TTAATGATGGAATAGCTTGGCTGTGATGAATCAGAAGCGGCGAGCCCTCTCTCGGCATGAAC	2260
QY	841	TCAGCGCGCTCCGCTGTGAGAGCTCTGGACCGTAG	873
Db	2261	TCAGCGCGCTCCGCTGTGAGAGCTCTGGACCGTAG	2293

<hr/>					
RESULT 2					
US-09-738-626-1/c					
; Sequence 1, Application US/09738626					
; Publication No. US20020197605A1					
; GENERAL INFORMATION:					
APPLICANT: NAKAGAWA, SATOSHI					
APPLICANT: MIZOGUCHI, HIROSHI					
APPLICANT: ANDO, SEIKO					
APPLICANT: HAYASHI, MIKIRO					
APPLICANT: OCHIAI, KEIKO					
APPLICANT: YOKOI, HARUHIKO					
APPLICANT: TATEISHI, NAKO					
APPLICANT: SENOH, AKIHIRO					
APPLICANT: IKEDA, MASATO					
APPLICANT: OZAKI, AKIO					
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES					
FILE REFERENCE: 249-125					
CURRENT APPLICATION NUMBER: US/09/738, 626					
CURRENT FILING DATE: 2000-12-18					
<hr/>					
QY	541	TGCCCGGTTCTTGGCGCATGTCAATTCAGCTTGCGGACATAATCAATATTGTTCCAGTTC	600		
Dd	1328008	TGCCCGGTTCTTGGCGCATGTCAATTCAGCTTGCGGACCAATCAATATTGTTCCAGTTC	1327949		
QY	601	AACATGCCTCAGACAGGGACTTTACCTTGCTGGCGCGGAAACCCTCTGGAATTCATCG	660		
Dd	1327948	AACATGCCCTCAGACAGGGACTTTACCTTGCTGGCGCGGAAACCCTCTGGAATTCATCG	1327889		
QY	661	AGATATTGTTCCTGAGCAGGCGCTTCGGCAAGTGGTGAGAAAGCAATGACGCCAAGACCA	720		
Dd	1327888	AGATATTGTTCCTGAGCAGGCGCTTCGGCAAGTGGTGAGAAAGCAATGACGCCAAGACCA	1327829		
QY	721	TTGTTGGCAGCTGACTGCACAAGTTTCTCACCGTCATCGCCCGGTTCTCTCCACCCACGA	780		
Dd	1327828	TTGTTGGCAGCTGACTGCACAAGTTTCTCACCGTCATCGCCCGGTTCTCTCCACCCACGA	1327769		
QY	781	TTAATCATGGAATAGCTTCGGCTGATGAATCAGAAGCGGSCAGCCCTCTCCGCCATGAAC	840		
Dd	1327768	TTAATCATGGAATAGCTTCGGCTGATGAATCAGAAGCGGSCAGCCCTCTCCGCCATGAAC	1327709		

QY 841 TCAGCGGCTCGCTGTGAGCTCTGGACCGTAG 873
 Db 1327708 TCAGCGGCTCGCTGTGAGCTCTGGACCGTAG 1327676

RESULT 3

US-09-738-626-3454/c
 ; Sequence 3454, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIALI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: Patent in ver. 3.0
 ; SEQ ID NO 3454
 ; LENGTH: 627
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3454

Query Match 65.1%; Score 568; DB 9; Length 627;

Best Local Similarity 100.0%; Pred. No. 8.3e-174;
 Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 GTTTTCGGGGTTTGGAACTCGGTGCTCGCCCAAAATGTTGATCGCGGCTGCGGA 365
 Db 627 GTTTTCGGGGTTTGGAACTCGGTGCTCGCCCAAAATGTTGATCGCGGCTGCGGA 568
 QY 366 AATCTCATCGATGCGCTCCAACTCGCGGTGAGAAACTCCAAAGTTTGTGAGTGAATCAAG 425
 Db 567 AATCTCATCGATGCGCTCCAACTCGCGGTGAGAAACTCCAAAGTTTGTGAGTGAATCAAG 508
 QY 426 GCTGTTGTCAGCTGCTCACTGACGAGCACCACCAATGACACTGCTCAGGTATCCG 485
 Db 507 GCTGTTGTCAGCTGCTCACTGACGAGCACCACCAATGACACTGCTCAGGTATCCG 448
 QY 486 GCGTACTCTCTCTGCTCGGCGAGCACCACCAATGACACTGCTCAGGTATCCG 545
 Db 447 GCGTACTCTCTCTGCTCGGCGAGCACCACCAATGACACTGCTCAGGTATCCG 388
 QY 546 GCGTCTCTGCGGAGTGTATGAGTGTGCGGACCATATCAATATTTTTCAGTTCAACAT 605
 Db 387 GCGTCTCTGCGGAGTGTATGAGTGTGCGGACCATATCAATATTTTTCAGTTCAACAT 328
 QY 606 GCGTCTGAGCAGGACTTACCTGCTGCGGAGCACCCTCTGGAATTTCCATCGAGATA 665
 Db 327 GCGTCTGAGCAGGACTTACCTGCTGCGGAGCACCCTCTGGAATTTCCATCGAGATA 268
 QY 666 TTTTCTGCTGAGCAGGCTCGGCAAGTGTGAGAAAGCAATGAGCCCAAGACCATTTGT 725
 Db 267 TTTTCTGCTGAGCAGGCTCGGCAAGTGTGAGAAAGCAATGAGCCCAAGACCATTTGT 208
 QY 726 GGCAGGTGATGCAACAAGTTTCTCACCGTTCATCGCCGGTTCTCTCACCCACGATTAAT 785

Db 207 GGCAGCTGATGCAACAAGTTTCTCACCGTTCATCGCCGGTTCTCTCACCCACGATTAAT 148
 QY 786 GATGGAATAGCTTGCTGATGATCAAGAGCGGCGAGCGGCTCTCGCCATGAACTCAGC 845
 Db 147 GATGGAATAGCTTGCTGATGATCAAGAGCGGCGAGCGGCTCTCTCGCCATGAACTCAGC 88
 QY 846 GGCCTCGCTGTGAGCTCTCTGAGCCGTAG 873
 Db 87 GGCCTCGCTGTGAGCTCTCTGAGCCGTAG 60

RESULT 4

US-10-627-476-669/c
 ; Sequence 669, Application US/10627476
 ; Publication No. US20040030116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompejus, Mark
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schoder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Haberhauer, Gregor
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 ; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
 ; FILE REFERENCE: BGI-125CPCN
 ; CURRENT APPLICATION NUMBER: US/10/627,476
 ; CURRENT FILING DATE: 2003-07-25
 ; PRIOR APPLICATION NUMBER: 09/602,787
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: USN 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: DE 19931454.3
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931478.0
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931563.9
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19932122.1
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932124.8
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932125.6
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932128.0
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932180.9
 ; PRIOR FILING DATE: 1999-07-09
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 678
 ; SEQ ID NO 669
 ; LENGTH: 993
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(970)
 ; OTHER INFORMATION: RXN03164
 US-10-627-476-669

Query Match 46.3%; Score 404.2; DB 12; Length 993;

Best Local Similarity 98.8%; Pred. No. 1.7e-120;
 Matches 418; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 453 AGCACCACCAATCAATGACATGCTGCTACCGTA--TCGCGCGCTACTCTCTTGTCTCGCGCAGC 510
 Db 993 AGCACCACCAATCAATGACATGCTGCTACCGTAATCGCGCGCTACTCTCTTGTCTCGCGCAGC 934
 QY 511 ACCCATCAAGCGCCATCTGCGCAAGTGAATGCGCGGTTCTCGGCGGATGTAATTCAGC 570
 Db 933 ACCCATCAAGCGCCATCTGCGCAAGTGAATGCGCGGTTCTCGGCGGATGTAATTCAGC 874
 QY 571 TTGCGGACCAATCAATGATTTTTCAGTTCAACATGCGCTCAGAGGACCTTACCTCG 630

Db 873 TTGGGACCATATATATTTTCAAGTTCACATGCGCCCTCAGACAGGACTTACCCCTGG 814
 QY 631 CTGGCGGGGAAACCTCTCGAATTCATCGAGATATTTTCCGTGAGAGGCCCTTGGCA 690
 Db 813 CTGGCGGGGAAACCTCTCGAATTCATCGAGATATTTTCCGTGAGAGGCCCTTGGCA 754
 QY 691 AGTGTGTGAGAAAGCAATGACGCAAGACCAATTTTGGCAGCTGACTGCAACAAAGTTCTCA 750
 Db 753 AGTGTGTGAGAAAGCAATGACGCAAGACCAATTTTGGCAGCTGACTGCAACAAAGTTCTCA 694
 QY 751 CCGTCATCGCCCGGTTCTTCCACCCCAAGCAATTAATGATGGAATAGCTTGGCTGATGAATC 810
 Db 693 CCGTCATCGCCCGGTTCTTCCACCCCAAGCAATTAATGATGGAATAGCTTGGCTGATGAATC 694
 QY 811 AGAAGCGGGAGCGCCCTCTCCGCCCATGAACTCAGCGGCTCCGCTGTGAGCTCTGGACCG 870
 Db 633 AGAAGCGGGAGCGCCCTCTCCGCCCATGAACTCAGCGGCTCCGCTGTGAGCTCTGGACCG 574
 QY 871 TAG 873
 Db 573 TAG 571

RESULT 5
 US-09-746-660A-51
 ; Sequence 51, Application US/09746660A
 ; Publication No. US20030049804A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompeius, Markus
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schroder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Haberhauer, Gregor
 ; APPLICANT: Kim, Jun-Won
 ; APPLICANT: Lee, Heung-Schick
 ; APPLICANT: Hwang, Byung-Joon
 ; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING
 ; FILE OF INVENTION: METABOLIC PATHWAY PROTEINS
 ; FILE REFERENCE: BGI-121CP2
 ; CURRENT APPLICATION NUMBER: US/09/746,660A
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 09/606740
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 09/603124
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/142101
 ; PRIOR FILING DATE: 1999-07-02
 ; PRIOR APPLICATION NUMBER: 60/148613
 ; PRIOR FILING DATE: 1999-08-12
 ; PRIOR APPLICATION NUMBER: 60/187970
 ; PRIOR FILING DATE: 2000-03-09
 ; PRIOR APPLICATION NUMBER: DE 19931420.9
 ; PRIOR FILING DATE: 1999-07-08
 ; NUMBER OF SEQ ID NOS: 125
 ; SOFTWARE: PatentIn Vers. 2.0
 ; SEQ ID NO 51
 ; LENGTH: 822
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(799)
 ; OTHER INFORMATION: RXA01394
 US-09-746-660A-51

Query Match 37.3%; Score 326; DB 10; Length 822;
 Best Local Similarity 100.0%; Pred. No. 3.9e-95;
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCGATAAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGCGCAATCGCTGACCTGGTTG 60

Db 497 GTCGATAAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGCGCAATCGCTGACCTGGTTG 556
 QY 61 AACCCGAATCGGTATTTTGGACGCGTTTGTGTTTATCGGCGGCTCGCGGCAATACGCGC 120
 Db 557 AACCCGAATCGGTATTTTGGACGCGTTTGTGTTTATCGGCGGCTCGCGGCAATACGCGC 616
 QY 121 GACACCGGACGCGTGGATTTTTCGCCGCTGGGCGCTTCGCGGCAAGCCTGATCTGTTCCCG 180
 Db 617 GACACCGGACGCGTGGATTTTTCGCCGCTGGGCGCTTCGCGGCAAGCCTGATCTGTTCCCG 676
 QY 181 CTGGTGGGTTTCGCCGACGACGATTTGACGCCGCTGTCCAGCCCCAAGGTGTGGCGC 240
 Db 677 CTGGTGGGTTTCGCCGACGACGATTTGACGCCGCTGTCCAGCCCCAAGGTGTGGCGC 736
 QY 241 TGGATCAACGCTCGTTCGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
 Db 737 TGGATCAACGCTCGTTCGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 796
 QY 301 GGTTAGTTTTCGCCGGGTTTGGAAATC 326
 Db 797 GGTTAGTTTTCGCCGGGTTTGGAAATC 822

RESULT 6
 US-10-166-142-7
 ; Sequence 7, Application US/10166142
 ; Publication No. US20030124687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GUNJI, YOSHIYA
 ; APPLICANT: YASUEDA, HISASHI
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
 ; FILE OF INVENTION: ASSIMILATING BACTERIUM
 ; FILE REFERENCE: 223789US
 ; CURRENT APPLICATION NUMBER: US/10/166,142
 ; CURRENT FILING DATE: 2002-06-11
 ; PRIOR APPLICATION NUMBER: JP 2001-1777075
 ; PRIOR FILING DATE: 2001-06-12
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 711
 ; TYPE: DNA
 ; ORGANISM: Brevibacterium lactofermentum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(711)
 ; OTHER INFORMATION:
 US-10-166-142-7

Query Match 35.1%; Score 306; DB 14; Length 711;
 Best Local Similarity 100.0%; Pred. No. 1.1e-88;
 Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCGATAAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGCGCAATCGCTGACCTGGTTG 60
 Db 406 GTCGATAAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGCGCAATCGCTGACCTGGTTG 465
 QY 61 AACCCGAATCGGTATTTTGGACGCGTTTGTGTTTATCGGCGGCTCGCGGCAATACGCGC 120
 Db 466 AACCCGAATCGGTATTTTGGACGCGTTTGTGTTTATCGGCGGCTCGCGGCAATACGCGC 525
 QY 121 GACACCGGACGCGTGGATTTTTCGCCGCTGGGCGCTTCGCGGCAAGCCTGATCTGTTCCCG 180
 Db 526 GACACCGGACGCGTGGATTTTTCGCCGCTGGGCGCTTCGCGGCAAGCCTGATCTGTTCCCG 585
 QY 181 CTGGTGGGTTTCGCCGCGCAGCAGCATTGTTCAGCCCGCTGTCCAGCCCCAAGGTGTGGCGC 240
 Db 586 CTGGTGGGTTTCGCCGCGCAGCAGCATTGTTCAGCCCGCTGTCCAGCCCCAAGGTGTGGCGC 645
 QY 241 TGGATCAACGCTCGTTCGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 300
 Db 646 TGGATCAACGCTCGTTCGTGGCAGTTGTGATGACCGCATTTGGCCATCAAACTGATGTTGATG 705

QY 301 GGTAG 306
Db 706 GGTAG 711

RESULT 7

US-10-166-142-9
; Sequence 9, Application US/10166142
; Publication No. US20030124687A1
; GENERAL INFORMATION:
; APPLICANT: GUNJI, YOSHIYA
; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
; TITLE OF INVENTION: ASSIMILATING BACTERIUM
; FILE REFERENCE: 223789US
; CURRENT APPLICATION NUMBER: US/10/166,142
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001-1777075
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 712
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(375)
; OTHER INFORMATION:
US-10-166-142-9

Query Match 34.9%; Score 304.4; DB 14; Length 712;

Best Local Similarity 99.7%; Pred. No. 3.8e-88;
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTCGATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGCACTGGTTG 60
Db 407 GTCGATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGCACTGGTTG 466
QY 61 AACCCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCGTGGCGGCGCAATACGGC 120
Db 467 ARCCCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCGTGGCGGCGCAATACGGC 526
QY 121 GACACCGGACGGTGGATTTTCGCGCGTGGCGCGTGGCGCGCAAGCTGATCTGGTTCCCG 180
Db 527 GACACCGGACGGTGGATTTTCGCGCGTGGCGCGTGGCGCGCAAGCTGATCTGGTTCCCG 586
QY 181 CTGGTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGTGTCCAGCCCCCAAGGTGTGGCGC 240
Db 587 CTGGTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGTGTCCAGCCCCCAAGGTGTGGCGC 646
QY 241 TGGATCAACGTCGTGGCGAGTGTGATGACCCGATGGCCATCAAACTGATGTTGATG 300
Db 647 TGGATCAACGTCGTGGCGAGTGTGATGACCCGATGGCCATCAAACTGATGTTGATG 706
QY 301 GGTAG 306
Db 707 GGTAG 712

RESULT 8

US-09-738-626-3455
; Sequence 3455, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOHI, AKIHIRO

; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3455
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3455

Query Match 34.7%; Score 303; DB 9; Length 708;

Best Local Similarity 100.0%; Pred. No. 1.1e-87;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGCACTGGTTG 60
Db 406 GTCGATAGCAGCGGGTTTGGGTAAAGCCCATGTTGATGGCAATCGTGTGCACTGGTTG 465
QY 61 AACCCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCGTGGCGGCGCAATACGGC 120
Db 466 AACCCGAATGCGTATTTGGACGCGTTTGTGTTATCGCGCGTGGCGGCGCAATACGGC 525
QY 121 GACACCGGACGGTGGATTTTCGCGCGTGGCGCGTGGCGCGCAAGCTGATCTGGTTCCCG 180
Db 526 GACACCGGACGGTGGATTTTCGCGCGTGGCGCGTGGCGCGCAAGCTGATCTGGTTCCCG 585
QY 181 CTGGTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGTGTCCAGCCCCCAAGGTGTGGCGC 240
Db 586 CTGGTGGGTTTCGCGCGCAGCAGCATTTGTCACGCCCGTGTCCAGCCCCCAAGGTGTGGCGC 645
QY 241 TGGATCAACGTCGTGGCGAGTGTGATGACCCGATGGCCATCAAACTGATGTTGATG 300
Db 646 TGGATCAACGTCGTGGCGAGTGTGATGACCCGATGGCCATCAAACTGATGTTGATG 705
QY 301 GGT 303
Db 706 GGT 708

RESULT 9

US-10-282-122A-17521/c
; Sequence 17521, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 17521
;; LENGTH: 1095
;; TYPE: DNA
;; ORGANISM: *Corynebacterium diptheriae*
US-10-282-122A-17521

Query Match 29.0%; Score 253.6; DB 12; Length 1095;
Best Local Similarity 65.6%; Pred. No. 1.5e-71;
Matches 370; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 310 TCGCGGCTTTTGGAAATCGGTGCTTTCGCCCAATGTTGATCGCGCGCTCGTGGGAAATC 369
DB 1094 TCGTGAACCTTCAGGCGAGTAGACACCGGCCAAATTAATTCACGCTCTTGGCAACA 1035

QY 370 TCATCGATCGCCTCCAACTCGGCGTCAGAAAATCTCAAGTTGTTGAGTGAATCAAGCTG 429
DB 1034 GTGTCAATATAATTCGTTTCCTCAACGCTGAATCAAGATTAATTAAGAGCTCCCAATTT 975

QY 430 TTGTCCAGCTGCTCAACTGAGGAGCAGCAACCAATCAATGCTGCTGACGCTATCCGCGCG 489
DB 974 TGATCCAACTGCGCTACCGAGATGCTCCGATCAACGCACTGGTTACCGTTTGGCCCCA 915

QY 490 TACTCTCTCTGCTCGCGCAGCACCATGCAAGCGCATCTGCGCAAGTCACTGCCCCTG 549
DB 914 TAAATGCTTGTTCGCGCAGCACCACCAACCAATAGCCAGCTGTGCAACGTTTGGCCAGT 855

QY 550 TCCTGGGCGATGTCATGAGTGTGGGACCAATATCAATATTTGTTCAAGTTCAAGTCAAGTCC 609
DB 854 CGTAAACGCTAAATCATTCAGTGCAGCAGCACCATGTCAGATTTTTCAGCTGATCTCT 795

QY 610 TCAGACAGGACTTACCTGCTGCGCGGAGCACTCTGGAATTCATCGAGATATTG 669
DB 794 TTGCCCAAGGACTTCCCTGCTGCTGCGCGTGAATCAATGCGACGCTCGAGATACCTG 735

QY 670 TCCGTGACGAGGCGCTCGCGCAAGTGTGAGAAAGCAATGACGCCAAGACCATTTGTGCA 729
DB 734 TCAGTTAAAGCCCTGTCAAGTGGGGAATGCAATAACACCAAGTCCGTTATTCGCC 675

QY 730 GCTGACTGCAACAGTTCTCAGCGTCAATCCCGGTTCTCCACCCCAAGATTAATGATG 789
DB 674 GCAGACTCAACAGGTTCTCGCCATCTTCACTGTTCTTACCCAGCGGTTCAAGATA 615

QY 790 GAATAGCTTGGCTGATGAATCAGAAAGCGGCGAGCCCTCTCGCCATGAATCAAGCGCC 849
DB 614 GAATAGCTCGGTTGATGAATCAGAGCGGCGAGCCCTCACCAGCTAGATTTCCAGATC 555

QY 850 TCCGCTGTGAGCTCTGACCCGTAG 873
DB 554 TCTCGGTTAGTTCAGGTCCTCATAG 531

RESULT 10
US-10-369-493-35374/c

;; Sequence 35374, Application US/10369493
;; Publication No. US20030233675A1
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Chen, Xianfeng
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; CURRENT FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 35374
;; LENGTH: 1026
;; TYPE: DNA
;; ORGANISM: *Agrobacterium tumefaciens*
US-10-369-493-35374

Query Match 15.6%; Score 136.6; DB 15; Length 1026;
Best Local Similarity 56.8%; Pred. No. 1.4e-33;
Matches 312; Conservative 0; Mismatches 204; Indels 33; Gaps 2;

QY 325 TCGGTGCGCTTCGCCCAATGTTGATCGCGCGTCTGCGGAAATCTCATCATCGCTCC 384
DB 1022 TCGGCAATTTGCCGAGGTTGATATCGCATCTTGGCGTAAACGGTCAATTCGGCC 963

QY 385 AACTCGGCGTCAGAAATCTCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGTCTCA 444
DB 962 AGCTCTCGGTAGAGAACTCGGCATTTATGAGTGTCTTTCACGAGTCTCTCGACCTGTTG 903

QY 445 ACTGACGAGCACCAATCAATGCTGCTGCTGACGATATCGCGCGTACTCTCTTCTGCTG 504
DB 902 ACAGGCTTGGCCCAATCAATGCTGAGGTAATGCGCGCGC-----CG 861

QY 505 CGCAGCAGCCATCAAGCGCCATCTGCGCAAGTGACTGCCCGGTTCTCGGCGAGATGCA 564
DB 860 CGCAGACCCAGGCAATTCGCATCTGCGCGAGGCTGCGCCAGCGCTCGGCAATGCTG 801

QY 565 TTGAGCTTGGGCAATATCAATATTTGTTCAATGTTTCAATGCTGCAATGCTGCAAGGACTTA 624
DB 800 TTCAAGCGCGGATGTTTTCAGCATTTGCTGTTGAGAAAGCGCGGTTGAGTGAATG 741

QY 625 CCCTGCTGCGCGGCAACCTCTGGAATTCATGAGATATTTGTCGTTGAGCAGGCGCC 684
DB 740 CTCTGTGAGGCAAGGCTGCTGCTGCGCACACCGCCAGATATTTGCTGCTCAGCATGCC 681

QY 685 TCGCAAGTGTGAGAAAGCAATGACGCGCAAGCAATTTGTTGGCAGCTGACTGCAACAG 744
DB 680 TCGCCAGCGGCAAGGAGATGAGCGCATACC-----CAGTCTTCC 636

QY 745 TTCTCAGCGTCAATCGCGGTTCTTCCACCAACGATTAATGATGAATGATGCTTGGCTGA 804
DB 635 AGCGTATCGAAGACCGCTTCTTCCATCCAGCGGTTGATCATGAAATGCTCGGCTGG 576

QY 805 TGAATCAGAGCGGCGAGCCCTCTTCCGCAATGAATCAGCGCGCTCCGCTGTGAGCTCT 864
DB 575 TGATGATGAGGCGGCTGCGAGATCTTTCAGGATAGCGGCGCTCGCGGTTGCGCTTC 516

QY 865 GCAGCGTAG 873
DB 515 GAGTTGTAG 507

RESULT 11
US-10-369-493-38826/c
;; Sequence 38826, Application US/10369493
;; Publication No. US20030233675A1
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2002-02-28
 ; NUMBER OF SEQ ID NOS: 47374
 ; TYPE: DNA
 ; ORGANISM: Agrobacterium tumefaciens
 ; US-10-369-493-38826
 ; LENGTH: 1026
 ; TYPE: DNA
 ; ORGANISM: Agrobacterium tumefaciens
 ; US-10-369-493-38826

Query Match 15.6%; Score 136.6; DB 15; Length 1026;
 Best Local Similarity 56.8%; Pred. No. 1.4e-33;
 Matches 312; Conservative 0; Mismatches 204; Indels 33; Gaps 2;
 QY 325 TCGTGGCCCTTCGCCCAATGTTGATCCCGGCTGCTGGGAAATCTCATCGATCGCTCC 384
 Db 1022 TCGGCAGATTTTGGCCAGAGTTGATATCCGCATCTTGGCGTAAACGGTTCGATTCGGCC 963
 QY 385 AACTCGGCGTCAGAAACTCCCAAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGTCTCA 444
 Db 962 AGCTCTCGGTAGAGACTTCGGCATTCAGTGTCTTCAGCAGTCTTCGACCTGTTCG 903
 QY 445 ACTGAGAGCAACCAATCAATGCACTGTCACGCTATCCCGCGCTACTCTCTCTTGTCTCG 504
 Db 902 ACACGGCTTGGCCCAATCAATGCTGAGTAAATGCGGCGC-----CG 861
 QY 505 CGCAGACCCATGCAAGCGCCATCTGCGCAAGTGAATGCTGCGGCTTCTCGGCGATGTCA 564
 Db 860 CGCAGAACCCAGGCAATTTGCCATCTCGCCAGCTGTGGCCACCGCGCTCGGCAATGCTG 801
 QY 565 TTGAGCTTGGGACCATATCAATATTGTTCACTTCAACATGCTTCAACATGCTTCAACATGCTTCA 624
 Db 800 TTGAGCGCGGAGTGTTCGACATTCGCTGCTGTTGAGAGCGCGGTTGAGTGTCTG 741
 QY 625 CCCTGGCTGGCGGGAAACCTCTGGAATTCATCGAGATATTGTTGCTGAGCAGGCGCC 684
 Db 740 CTCTGTGAGGCACTGCTCCATCCGACACCGCCAGATATTTCGTGCTCAGCATGCC 681
 QY 685 TCGCAGTGTGAGAGCAATGACCGCAAGCAATGTTGTCAGCTGATCTGCAACAG 744
 Db 680 TCGCCAGCGCGGAGAGAGCATGGAGCGGATACC-----CAGTTCTTCC 636
 QY 745 TTCTCACCGTCATCGCCGGTTCTTCCACCCAAAGATTAATGATGGAATAGCTTGGCTGA 804
 Db 635 AGCGTATCGACAGACCGTCTTCTCGATCCAGCGTGTGATCATCGATAGCTCGGCTGG 576
 QY 805 TGAATCAGAGCGGCGCCCTCTCTCGCATGAACTCAGCGGCTTCGCTGTGAGCTCT 864
 Db 575 TGGATGATGAGGCGGCTGCGAGATCTTCAGGATAGCGGCGCTTCGCGGCTCGGCTTC 516
 QY 865 GGACCGTAG 873
 Db 515 GAGTTGTAG 507

RESULT 12
 US-10-369-493-38124/c
 ; Sequence 38124, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 38124
 ; LENGTH: 1028
 ; TYPE: DNA
 ; ORGANISM: Agrobacterium tumefaciens
 ; US-10-369-493-38124

Query Match 15.6%; Score 136.6; DB 15; Length 1028;
 Best Local Similarity 56.8%; Pred. No. 1.5e-33;
 Matches 312; Conservative 0; Mismatches 204; Indels 33; Gaps 2;
 QY 325 TCGTGGCCCTTCGCCCAATGTTGATCCCGGCTGCTGGGAAATCTCATCGATCGCTCC 384
 Db 1021 TCGGCAGATTTTGGCCAGAGTTGATATCCGCATCTTGGCGTAAACGGTTCGATTCGGCC 962
 QY 385 AACTCGGCGTCAGAAACTCCCAAGTTGTTGAGTGAATCAAGGCTTGTTCAGCTGTCTCA 444
 Db 961 AGCTCTCGGTAGAGACTTCGGCATTCAGTGTCTTCAGCAGTCTTCGACCTGTTCG 902
 QY 445 ACTGAGAGCAACCAATCAATGCACTGTCACGCTATCCCGCGCTACTCTCTCTTGTCTCG 504
 Db 901 ACACGGCTTGGCCCAATCAATGCTGAGTAAATGCGGCGC-----CG 860
 QY 505 CGCAGACCCATGCAAGCGCCATCTGCGCAAGTGAATGCTGCGGCTTCTCGGCGATGTCA 564
 Db 859 CGCAGAACCCAGGCAATTTGCCATCTCGCCAGCTGTGGCCACCGCGCTCGGCAATGCTG 800
 QY 565 TTGAGCTTGGGACCATATCAATATTGTTCACTTCAACATGCTTCAACATGCTTCAACATGCTTCA 624
 Db 799 TTGAGCGCGGAGTGTTCGACATTCGCTGCTGTTGAGAGCGCGGTTGAGTGTCTG 740
 QY 625 CCCTGGCTGGCGGGAAACCTCTGGAATTCATCGAGATATTGTTGCTGAGCAGGCGCC 684
 Db 739 CTCTGTGAGGCACTGCTCCATCCGACACCGCCAGATATTTCGTGCTCAGCATGCC 680
 QY 685 TCGCAGTGTGAGAGCAATGACCGCAAGCAATGTTGTTGAGCAGTGTCTGCAACAG 744
 Db 679 TCGCCAGCGCGGAGAGAGCATGGAGCGGATACC-----CAGTTCTTCC 635
 QY 745 TTCTCACCGTCATCGCCGGTTCTTCCACCCAAAGATTAATGATGGAATAGCTTGGCTGA 804
 Db 634 AGCGTATCGACAGACCGTCTTCTCGATCCAGCGTGTGATCATCGATAGCTCGGCTGG 575
 QY 805 TGAATCAGAGCGGCGCCCTCTCTCGCATGAACTCAGCGGCTTCGCTGTGAGCTCT 864
 Db 574 TGGATGATGAGGCGGCTGCGAGATCTTCAGGATAGCGGCGCTTCGCGGCTCGGCTTC 515
 QY 865 GGACCGTAG 873
 Db 514 GAGTTGTAG 506

RESULT 13
 US-10-282-122A-13437/c
 ; Sequence 13437, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant

```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13437
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Burkholderia fungorum
US-10-282-122A-13437

Query Match      12.1%; Score 105.8; DB 12; Length 1041;
Best Local Similarity 55.6%; Pred. No. 1.5e-23;
Matches 293; Conservative 0; Mismatches 192; Indels 42; Gaps 3;

QY 331 GCCTTCGGCCAAATGTTGATCGCGGCTCTGGGAAATCTCATCGATCGCTCCCACTCG 390
DB 1022 GGCCTTTTCCACAGGTTGATCGCGCTTCCGTCGCGTAGCGGTGATCTCGCGCAGCTCG 963

QY 391 GCGTCAGAAATCCAAAGTTGTGAGTGAATCAAGCTGTTGTCCAGCTGCTCAACTGAC 450
DB 962 TCCCTTCGAGAACCGAAGTTCTTCAGCGCCGACGTTTTCAGCACCTGCTCCGCGCG 903

QY 451 GAAGCACAATCAATCACTGTCACCGGTATCCGCGCGGTACTCTCTCTGCTCGCGCAGC 510
DB 902 CTCGCGCGATCAGCACGGAAGTCACG-----CGCGGATCGCGCAGC 861

QY 511 ACCATGCAAGCGCGATCTGCGCAAGTGAAGTCTGCGCGCTTCTCGCGGATGATTCATGAGC 570
DB 860 GCCACGCGAGCGCGATTTGTGCGAGGCTCTGACGCGACGTTGCGGATGCTGTTGAGC 801

QY 571 TTGCGGACCATATCAATATTGTTCAAGTTTCAATGCCCTCAGACGAGGACTTACC----- 626
DB 800 TTGCGGACGCTCTGATATTCTGCGGCTCAGATGCTCTGTTCAACGAACCGCGCGCC 741

QY 627 --CTGGCTGGCGCGGGAACCCCTCTGGAATTCATCAGATATTTGTCGAGAGAGGCGCC 684
DB 740 GGCCTTTGTCGCGCGCTCTTCCGCGACGCGCTTCAGATATTGCTCTGTCGAGAGCGCC 681

QY 685 TGCGCAAGTGTGAGAACGATGACGCGACGACCATTTGTCGAGCTGACTGCAACAAAG 744
DB 680 TGAGCAGCGCGGTGAACGCGATGCGCGCGCGCGCGCGCTTCTCAATGCGCGCGCAAA-- 622

QY 745 TTCTCACCGTCATCGCCCGGCTTCTTCCACCAAGATTAATGATGAATAGTCTTGCTGA 804
DB 621 -----CTCATGCTCTGATCCAGCATTTGATGAGAGGCTGA 579

QY 805 TGAATCAGAACGCGGCGAGCGCTCTCTCCGCCATGAATCTCAGCGCGCTC 851

```

```

DB 578 TGAATCAGACGCGGACCTTGTATTTCGGGACGACGCTTCGCCATTTC 532
|||||
RESULT 14
US-10-282-122A-14390/c
; Sequence 14390, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14390
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-14390

Query Match      11.1%; Score 96.8; DB 12; Length 1107;
Best Local Similarity 56.9%; Pred. No. 1.3e-20;
Matches 232; Conservative 0; Mismatches 152; Indels 24; Gaps 2;

QY 327 GGTGGCTTCGCCCAATGTTGATCGCGGCTCTGGGAAATCTCATCGATCGCTCCAA 386
DB 1092 GGAGGCTTTTCCCAAGATTTGATCCGCCCTTCGGTCGCATAGCGGTGATCTCGCGCAT 1033

QY 387 CTCGGGCTCAGAAAATCTCCAGTTGTTGAGTGAATCAAGCTGTTGTCAGCTGCTCAAC 446
DB 1032 CTCGTCGTCTGTAACCGAGATGGCGCGCGCGAGCTTCTCGCGCACCTGCTTCGCG 973

QY 447 TGACGAAGCAACCAATCAATGTCAGCTGTCAGCGTATCCGCGCGTACTCTCTTGTCTCGCG 506
DB 972 GCGGCTCGCACCGATCAACGCGGAGTCAAG-----CGCGAATCGCG 931

QY 507 CAGCACCATTCAAGCGCATCTCGCGCAAGTGAATCGCCCGGCTTCTCGCGCGATGTCATT 566
DB 930 CAGCACCAGGAGCGCATCTCGCGGAGGCTCTCGCGCGCGCGCTCTCGCGCATCTCGTT 871

```

QY 567 GAGCTTGGGACCATATCAATATTGTTACGTTCAACATGCGCTCAGACAGGACTTACC 626
 Db 870 CAGCTTGGCGCAGTCTCGAGGTTCTCCGCGCTCAGGTGAGTCTTCTCAGCATCGCC 811
 QY 627 -----CTGGCTGGCGGGAAACCTCTGGAATTCATCGATATTTGTCGGTGAAG 680
 Db 810 GCCGGCTTTGTTGATCGCGCATCCGCGGCACGCGTTTCAGATATCTTCAGAGTCAAG 751
 QY 691 GCCCTGGCAAGTGTGAGAAAGCAATGACGCCAGACCAATTGTTGC 728
 Db 750 CCGCTGGGAGCGCGGTGAACGCGATGACGCCGCGCTGCTGTC 703

RESULT 15

US-10-282-122A-11887/c
 ; Sequence 11887, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangou
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 11887
 ; LENGTH: 1041
 ; TYPE: DNA
 ; ORGANISM: Burkholderia cepacia
 US-10-282-122A-11887

Query Match 11.0%; Score 96.2; DB 12; Length 1041;
 Best Local Similarity 55.7%; Pred. No. 1.9e-20;
 Matches 241; Conservative 0; Mismatches 168; Indels 24; Gaps 2;
 QY 327 GGTGGCTTGGCCAAATGTTGATGCGCGGTCTGGGAAATCTCATCGATCGCTCCAA 386
 Db 1026 GGACGGCTTTTCCACAGATTGATGCGCGCTTCCGTCGCGTAAGATCGATCTCGCGAG 967
 QY 387 CTCGGCGTCAGAAACTCCAAAGTTGTTGAGTGAATCAAGGCTGTGTCCAGCTGCTCAAC 446

Db 966 TTCTCCGCGGAGAAATTCAGGTTCTGCAACGACCGACGTTCTCGCGCACCTGTCTGGC 907
 QY 447 TGAAGAGACCAATCAATGACATGCTGCTCAAGTATCCGCGCGTACTCTCTCTTGTCTGG 506
 Db 906 CCGGCTCGCGCGGATCAGCGGAGGTTCACGCGCG-----TTACG 865
 QY 507 CAGCACCCATGCAAGCGGCATCTGCGCAAGTGAATGCGCGGTTCTTGGGCGATGCTATT 566
 Db 864 CAGCACCCATGCAAGCGGCATCTGCGCAAGTGAATGCGCGGTTCTTGGGCGATGCTATT 805
 QY 567 GAGCTTGGGACCATATCAATATTGTTTCAAGTTCACATGCGCTCAGACAGGACTTACC 626
 Db 804 GAGCTTGGGACCATATCAATATTGTTTCAAGTTCACATGCGCTCAGACAGGACTTACC 745
 QY 627 -----CTGGCTGGCGCGGAAACCTCTGGAATTCATCGAGATATTTGTCGGTGAAG 680
 Db 744 GCCCGGTTTGTTCACGCGCGGTCTCGCGCGGCGGTTTCAAGTACTTTCAGTGAAGCAA 685
 QY 681 GCCCTGGCGCAAGTGTGAGAAAGCAATGACCGCAAGACCATTTGTTGGCAGCTGACTGCAA 740
 Db 684 GCCCTGGCGCGGCGGTGAACGCGATGCTGCGCGCGGCGGCGTCTGTCGAGCGTACCGAG 625
 QY 741 CAGCTTCTCACCG 753
 Db 624 CAGATCTCTCTCG 612

Search completed: March 17, 2004, 09:47:12
 Job time : 326.967 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:43:00 ; Search time 2400.75 Seconds
(without alignments)
10858.975 Million cell updates/sec

Title: US-09-105-117K-1_COPY_1421_2293

Perfect score: 873

Sequence: 1 gfcgataagcagcgggttg.....ctgtgagctctggaccgttag 873

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em estba:**

2: em esthum:**

3: em estin:**

4: em estmu:**

5: em estov:**

6: em estpl:**

7: em estro:**

8: em htc:**

9: gb est1:**

10: gb est2:**

11: gb htc:**

12: gb est3:**

13: gb est4:**

14: gb est5:**

15: em estfun:**

16: em estom:**

17: em ges hum:**

18: em ges inv:**

19: em ges pln:**

20: em ges vrt:**

21: em ges fun:**

22: em ges man:**

23: em ges mus:**

24: em ges pro:**

25: em ges rod:**

26: em ges phg:**

27: em ges vri:**

28: gb gss1:**

29: gb gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	56.4	6.5	860	28 AF075981	AF075981 AF075981
C 2	54.2	6.2	897	28 BZ560550	BZ560550 pacs2-164
C 3	51.4	5.9	1620	28 BZ568946	BZ568946 pacs2-164
C 4	48.8	5.6	1101	29 CNS0175Y	AL108460 Drosophil

5	45	5.2	720	14	CD881613
6	44.4	5.1	664	29	CNS03K4J
7	43.8	5.0	3237	28	BH770958
8	43.2	4.9	1395	28	BZ574094
9	42.4	4.9	699	13	BX424825
10	42	4.8	457	13	BY246241
11	41.8	4.8	538	12	CD861975
12	41.6	4.8	558	12	BM712036
13	41.6	4.8	588	13	BU220119
14	41	4.7	617	12	BM729055
15	41	4.7	770	10	BE898495
16	41	4.7	861	10	BE740880
17	40	4.6	357	9	AI418751
18	40	4.6	431	9	AA452858
19	40	4.6	447	14	W91005
20	40	4.6	463	9	AA476500
21	40	4.6	504	9	AI159901
22	40	4.6	537	29	CG694695
23	40	4.6	540	9	AI401605
24	40	4.6	564	9	AI820014
25	40	4.6	600	12	BG806166
26	40	4.6	619	14	CF533294
27	40	4.6	656	28	BZ894552
28	40	4.6	3760	11	AK031268
29	40	4.6	5069	29	AY405421
30	40	4.6	5094	29	AY405420
31	39.8	4.6	1101	29	CNS00KK2
32	39.6	4.5	925	29	CNS0091P
33	39.4	4.5	546	14	CF879615
34	39.2	4.5	846	29	CNS010RJ
35	39.2	4.5	1201	13	BX361080
36	39.2	4.5	1201	13	BX381961
37	39	4.5	543	14	CB640932
38	38.8	4.4	1201	13	BX381961
39	38.6	4.4	436	10	BE025218
40	38.6	4.4	513	13	BQ818660
41	38.6	4.4	828	29	CNS026XT
42	38.6	4.4	830	29	CNS01MQH
43	38.6	4.4	882	14	CF241805
44	38.6	4.4	1201	13	BX356664
45	38.6	4.4	1667	11	AY108197

ALIGNMENTS

RESULT 1	AF075981/c	860 bp	DNA	linear	GSS 29-AUG-2000
LOCUS	AF075981	Salmonella typhimurium LT2	Lambda DASH II	Salmonella	
DEFINITION	AF075981	Salmonella typhimurium genomic clone 390-T3	genomic survey	sequence	
ACCESSION	AF075981				
VERSION	AF075981.1	GI:3320851			
KEYWORDS	GSS				
SOURCE	Salmonella typhimurium				
ORGANISM	Salmonella typhimurium				
REFERENCE	1 (bases 1 to 860)				
AUTHORS	Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.				
TITLE	Sample sequencing of a Salmonella typhimurium LT2 lambda library:				
JOURNAL	comparison to the Escherichia coli K12 genome				
MEDLINE	FEMS Microbiol. Lett. 173 (2), 411-423 (1999)				
PUBMED	99243757				
COMMENT	10227170				
	Contact: McClelland M				
	Molecular Biology				
	Sidney Kimmel Cancer Center				
	3099 Science Park Road, San Diego, CA 92121, USA				
	Email: mclelland@lifsci.sdu.edu				
	Class: shotgun.				
	Location/Qualifiers				
	1. .860				

```
/organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="L12"
/db_xref="taxon:602"
/clone="390-T3"
/notes="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"

ORIGIN
Query Match 6.5%; Score 56.4; DB 28; Length 860;
Best Local Similarity 53.5%; Pred. No. 0.01;
Matches 199; Conservative 0; Mismatches 151; Indels 22; Gaps 3;
QY 505 CGCAGCACCCATCGAGCCATCTGCGCAAGTGAATCGCGGCTTCCTGGGCGATGTCA 564
Db 400 CGCAGCATCCAGCGAGCGCCATTTGATACATTTTGGACCGGTGCGGCGCAGTTC 341
QY 565 TTGAGCTTGGCGACCATATCATATTTCTTCACTTCAATCGCTTCAATCGCTTCAAGAGGACTTA 624
Db 340 TTCAACCGCGAATCTTTTCCAGTTTGTGCGGGTAACTGTCTGTGTTTAAGGAAACGG 281
QY 625 CCTCGCTG---GGCGGGAACCTCTGGAATTCATCGAATATTTGCGGTGCTGCGTGAAG 681
Db 280 CTTCCGCTCGCGCGCGGGAATCTTCCGGAATACCAATTCAAATTAACGTCAGTGAGC-TG 222
QY 682 CCTCGCAAGTGTGAGAAAGCAATGACGCAAGACCAATTTGGCAGCTGACTGCAAC 741
Db 221 CCGGCCCGCAGCGAGAGCAATACTACGACGCTTTTCTCTACACAGCCGACG 162
QY 742 AAGTCTTCAACGTCATCGCGGTTCTTCCACCAACAGTAATTAATGATGAATGCTTGGC 801
Db 161 AG-----CCGCTCTCTACCAACGCTCAAAAAGCGAATATTATAGC 120
QY 802 TGATGAATCAGACCGCGGAGCCCTCTCCGCGCATCACTCAGCGCTCCGCTGAGC 861
Db 119 TGATGAATCAGGCAAGCGTGGCGGATCTTCCAGAAATATCATAGCTTCTTGCCGAGA 60
QY 862 TCTGGACCGTAG 873
Db 59 TCGGACGATAG 48

RESULT 2
BZ560550/c
LOCUS BZ560550 897 bp DNA linear GSS 17-DEC-2002
DEFINITION pacs2-164_2569.xl pacs2-164 Pseudomonas aeruginosa genomic clone
pacs2-164_2569, genomic survey sequence.
ACCESSION BZ560550
VERSION BZ560550.1 GI:27178749
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 897)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) in press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
1..897
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="L12"
/db_xref="taxon:602"
/clone="390-T3"
/notes="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"
```

```
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_2569"
/clone_lib="pacs2-164"
/notes="clinical isolate 2-164 Whole genomic shotgun
library."

ORIGIN
Query Match 6.2%; Score 54.2; DB 28; Length 897;
Best Local Similarity 52.8%; Pred. No. 0.035;
Matches 143; Conservative 0; Mismatches 120; Indels 8; Gaps 1;
QY 35 TGATGGCAATCGTCTGACATCTGCTTGAACCCGGAATCGGTATTTGACGCGTTGTGTTA 94
Db 552 TGGGGGCACTCGGGCTTCACTGCTTAAACCCCAACGCTCTATTCGATACCGTACTTCTTAC 493
QY 95 TCGGGCGGCTCGGGCGCAATACGGCGACACCGGCTGGAATTTTCGCGCTTGGCGGT 154
Db 492 GGTTCGTCGGCGCGCAAGNCCGCG-----CCGGGCGCATATGCCCTCGGGCGG 441
QY 155 TCGCGCGAAGCTGATCTGCTTCCCGCTGCTGGGTTTGGCGCAGCAGCATTTGTACGCC 214
Db 440 CCAAGCGCTCGCTGATGTGTTCTTCGCGCTCGCCCTCGGCGCGCATGGCTGGCCCCCT 381
QY 215 CGCTGTCAGCCCCCAAGGTGTGGCGCTGGATCAACGCTGCTGCGCAGTTGTGATGACCG 274
Db 380 GGTCTGGCGCGCGCGGCGCACCTGGCGCTGCTCGACCTGATGTGGCGGCCATGATCTGG 321
QY 275 CATTGGCCATCAAACTGATGTTTGATGGGTTA 305
Db 320 GCATGGCGCGCACTGCTTTCGCGGGATA 290

RESULT 3
BZ568946/c
LOCUS BZ568946 1620 bp DNA linear GSS 17-DEC-2002
DEFINITION pacs2-164_8165.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
pacs2-164_8165, genomic survey sequence.
ACCESSION BZ568946
VERSION BZ568946.1 GI:27202770
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1620)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) in press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
1..1620
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_8165"
/clone_lib="pacs2-164"
/notes="clinical isolate 2-164 Whole genomic shotgun
library."

ORIGIN
Query Match 5.9%; Score 51.4; DB 28; Length 1620;
Best Local Similarity 54.3%; Pred. No. 0.22;
```

```

Matches 100; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 122 ACACGGAGCGTGAATTTTCGGCGTGGCGGTCGGGCAAGCTGATCTGTTCCCGC 181
Db 618 AGCCGCGCGGCGGCATATGCGCTCGCGCGGCGGCGGCTCGATGATGTTCTTCG 559
QY 182 TGTGTTGTTTCGGCGGAGCAGCATGTGTCAGCCCGGTGTCCAGCCCAAGGTGTGGCGCT 241
Db 558 CCCTCGCCTCGGCGGCGATGCTGCGCCCTCGCTGGCGGCGCCGNCACCTGGCGCC 499
QY 242 GATCAAGCTGCTGGCGAGTTGTGATGACCGCATGGCGCATCAAACTGATTTGATGG 301
Db 498 TGTCTGATGATGTTGGCGGCATGCTGCGGCATGCTGGGCATGCGCGCAACTGCTGTTCCGG 439
QY 302 GTTA 305
Db 438 GATA 435

RESULT 4
CNS017SY/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN3/L08 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108460
VERSION AL108460.1 GI:5628764
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEFH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN3/L08"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : SP6"

ORIGIN
Query Match 5.6%; Score 48.8; DB 29; Length 1101;
Best Local Similarity 14.5%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 47; Conservative 156; Mismatches 121;

QY 72 GTATTGGACGGCTTGTGTTATCGCGGCGTGGCGGCATATACGCCGACACCGGACG 131
Db 1079 KTTTBTBTSSSTSTSTSTBTBSSBSSBSSBSSBSSBSSBSSBSSBSSBSS 1020
QY 132 GTGATTTTCGGCTGGCGGTTCGGCGAAGCTGATCTGCTCCGCTGGTGGGTTT 191
Db 1019 SSTSBTBTSTSSSTTTTTTTTTTTTBTBTBSSBSSBSSBSSBSSBSSBSS 960
QY 192 CGGCGGACGACATGTCACGCCCGCTGTCCAGCCCAAGGTGTGGCGCTGATCAACGT 251
Db 959 BSTSSASBSSSSSSBSSSTSTBTBSSBSSBSSBSSBSSBSSBSSBSSBSSBT 900

```

```

QY 252 CGTCGGCAGTTGTGATGACCGCATTTGCCATCAAACTGATGTTGATGGGTAGTTTTC 311
Db 899 BTXSTSTSSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 840
QY 312 CGCGGTTTGGAAATCGGTGGCGCTTCGCCCAATGTTGATCCGGCGTGTGGGAAATCTC 371
Db 839 TSGSTBTSTBMSKBSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 780
QY 372 ATCGATCGCTTCAACTCGGCGTC 395
Db 779 GBCSTGGSCCCYCCOCTCTC 756

RESULT 5
CD881613
LOCUS
DEFINITION F1.103L23F010329 F1 Triticum aestivum cDNA clone F1103L23, mRNA
sequence.
ACCESSION CD881613.1 GI:32641286
VERSION CD881613
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 720)
AUTHORS Genoplatte.
TITLE Genoplatte, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplatte
Genoplatte
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplatte' (http://www.genoplatte.com
and http://genoplatte.info.infobiogen.fr).
FEATURES
source
1..720
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="F1103L23"
/tissue_type="leaf one"
/clone_lib="F1"

ORIGIN
Query Match 5.2%; Score 45; DB 14; Length 720;
Best Local Similarity 48.0%; Pred. No. 5.2;
Matches 129; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 1 GTGATTAAGACGGGTTTGGGTAAGCCCATGTTGATGCAATGTCGTGACTGCTGTTG 60
Db 252 GTGGAGGAGAAGCTGTGGCGCTCAAGCCCAAGAGCATGCACTTCGCGAGGCGCGCGT 311
QY 61 AACCCGAATGCTGATTTTGGAGCGGTTTGTGTTATTCGGCGGCTGCGGCGCAATACGCG 120
Db 312 CTGCGCGCTGCGCATCGAGACGCGCATGAGGGCTTCGAGAGGCGGGCTTTCGCGCGCG 371
QY 121 GACACCGGACGGTGAATTTTTCGCGGCTGCGCGGTCGCGGCAAGCTGATGTTCCCG 180
Db 372 AAGTCCATCTCTGCTCTCGCGCGCGCGCGGAGTGGGACCTCTGTCATCCAGCTGGCG 431
QY 181 CTGCTGGGTTTTCGCGCGGACGACGATTTGTCAAGCCGCTGTCCAGCCCAAGGTGTGCGCG 240
Db 432 AAGCAAGTTTACGGGCGCTGTGACGGTGGCGGCGCACGCCAGCACCCCGAAGCTGGAGCTC 491
QY 241 TGGATCAAGCTGCTGCTGGCAGTTGTGAT 269
Db 492 CTGAAAGACCTGGAGCGCAGCTGGCCAT 520

```

RESULT 6
CNS03K4J
LOCUS
DEFINITION
CNS03K4J 664 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone
032806 of library G from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION
AL247708
VERSION
AL247708.1 GI:7968720
KEYWORDS
GSS; genome survey sequence.
SOURCE
Tetraodon nigroviridis
ORGANISM
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.

REFERENCE
1
AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Queier,F.,
Saurin,W. and Weissenbach,J.
TITLE
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL
Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE
20296633
PubMed
10835645

REFERENCE
2
AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL
Genome Res. 10 (7), 939-949 (2000)
MEDLINE
20359837
PubMed
10899143

REFERENCE
3 (bases 1 to 664)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES
source
1..664
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="032806"
/clone_lib="G"
/note="Genoscope sequence ID : C0B6032DA03LP1-end : T7"

ORIGIN
Query Match 5.1%; Score 44.4; DB 29; Length 664;
Best Local Similarity 48.2%; Pred. No. 7;
Matches 120; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 181 CTGGTGGGTTTCGGCGAGCAGCATTTGTCAGCCCGCTGTCCAGCCCAAGGTGGCGC 240
Db 278 CAGGTCCGCTCTCCGCCGGAATGCGGGGCCCGCGGAGCGGCGCGGCGGCGGCGC 337
QY 241 TGATCAACGTCGTGGGAGTTGTGATGACCGCATGGCCATCAACTGATGTTGATG 300
Db 338 CGCGGNTCTTTTGTGCGGCGGTGGCGCGGACAGGCCCGCCGCGCTGNTGGTGG 397
QY 301 GGTAGTTTTCGGGGTTTGGAAATCGGTGGCCTTGCCCAAAATGTTGATGCCGGGTGC 360
Db 398 GCGAAGCATCCGGTGGTGGGASCCGTGGCTCTCGCACCGATGGAGGCCCTGGGGCG 457
QY 361 TGGGAAATCTCATGATCGCTCCAACTCGGCGTCCAGAAACTCCAGTTGTTGATGAA 420

Db 458 GAGGCTCCCCACAGGTGGARCTCCCGCGCGTGGAGGCTCCCGCGTGGTGGGTCTT 517
QY 421 TCAGGCTG 429
Db 518 CCCGCGCG 526

RESULT 7
BH770958 3237 bp DNA linear GSS 01-MAY-2002
LOCUS
DEFINITION
LIMGtag686 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, genomic survey sequence.
ACCESSION
BH770958
VERSION
BH770958.1 GI:20373915
KEYWORDS
GSS.
SOURCE
Lactococcus lactis subsp. cremoris
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE
1 (bases 1 to 3237)
AUTHORS
Boltin,A., Ehrlich,S.D. and Sorokin,A.
TITLE
Studies of genomes of dairy bacteria Lactococcus lactis
Sci. Aliments (2002) In press
JOURNAL
Contact: Sorokin A
COMMENT
Genetique Microbienne
INRA
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokin@jouy.inra.fr
best homologue in strain IL1403 is nadR (98%)
Class: shotgun
High quality sequence start: 30
High quality sequence stop: 3207.
Location/Qualifiers
1..3237
/organism="Lactococcus lactis subsp. cremoris"
/mol_type="genomic DNA"
/strain="MG1363"
/sub_species="cremoris"
/db_xref="taxon:1359"
/clone_lib="MG1363 Random Sequence Tag Library"
/note="Vector: pSGM2; Site 1: SmaI; Library of
chromosomal fragments of L.lactis strain MG1363 was
prepared by partial AluI digestion or by sonication."

FEATURES
source
1..3237
Query Match 5.0%; Score 43.8; DB 28; Length 3237;
Best Local Similarity 49.9%; Pred. No. 22;
Matches 170; Conservative 0; Mismatches 162; Indels 9; Gaps 2;
QY 385 AACTCGGCTCAGAAAATCCAGTTGTTGAGTGAATCAAGGCTGTTGTCAGCTGTCA 444
Db 1956 AATTCTGTCAGTAATCCAGAGTTTAAAGCTGTAACAGCTTGAACCTTTGCCAGCTCTT--TCC 2015
QY 445 ACTGACGAGCAGCAACATCAATGCTACGTTACGATCCGCGCGCTACTCTCTGCTG 504
Db 2016 GGACGACTTGGCCCAATCAAGCTGATGTAACAGCTTGAACCTTTGCCAGCTCTT--TCC 2072
QY 505 CGAGCAGCCCATGACAGCCGCTCTGCGCAAGTCACTGCCCGTTCCTGGCGATGTCA 564
Db 2073 CGTAAACCAAGCTTATGCACTTTGAGCTAGATTTGTCACAGCTTTGAGCTAAATCA 2132
QY 565 TTGAGCTCTGGGACCATATCAATATTGTTTACGTTTCAATGCTTCAATGCTTCAATGCTTCA 624
Db 2133 TTAAGGGCTTGAACCTGTTTCAATCTTTCTGTGTTAGACTGTATCATGCAAGTCGA 2192
QY 625 CCCTGGCTGGCGCG--GAAACCCCTCTGGAATTCATTCAGATATTGTCGAGAC 678
Db 2193 TATGAGGATCAGCATTCGCGAATTTTCAGGAATTCATTCGAAATATTTCTCTGTTAAG 2252
QY 679 AGGCCCTGGCGAAGTGGTGAAGAAGCAATGAGCCAGACC 719

Db 2253 AGTCCTTGATACAAAGGCTTAAGGCAATTTGTTCTATCCC 2293

RESULT 8 BX574094/c

LOCUS
DEFINITION
msh2.3508.y3 msh Pseudomonas aeruginosa genomic clone msh2_3508,
genomic survey sequence.

ACCESSION
BX574094

VERSION
BX574094.1 GI:2709155

KEYWORDS
GSS.

SOURCE
Pseudomonas aeruginosa

ORGANISM
Pseudomonas aeruginosa

REFERENCE
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

AUTHORS
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

TITLE
Burns,J.L., Kaul,R. and Olsen,M.V.

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

CONTACT: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

Location/Qualifiers

1..1395

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="M5H"

/db_xref="taxon:287"

/clone="msh2_3508"

/note="Environmental isolate. Whole genomic shotgun

library."

ORIGIN

Query Match 4.9%; Score 43.2; DB 28; Length 1395;

Best Local Similarity 55.3%; Pred. No. 20;

Matches 84; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 488 CGTACTCTCTGCTCGCGCAGCACCACCGAGCCATCTGCGCAAGTCACTGCCGCGC 547

DB 527 CGTCTCGGCTTGTCACCGAGCACCAGCGCGGAGCACCATCGGGGGGATGGGTAACCGA 468

QY 548 GTTCTGCGGCGATGTCATTGAGCTTGCGGACCATATCAATATTGTTCACTTCAACATGC 607

DB 467 GTTCTTGTAACAGGTAGCGAGCGCGCGACAGCAGCATCAGGACCGATCGAGCATG 408

QY 608 CTTGAGCAGGAGCTTACCTTGGCTGGCGCGG 639

DB 407 CGTTGTGACCGAGTAGGCTCCGATGCGCGAG 376

RESULT 9

BX424825/c

LOCUS

DEFINITION

BX424825 Homo sapiens PLACENTA Homo sapiens CDNA clone

EST.

ACCESSION

BX424825

VERSION

BX424825.1 GI:30647817

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 699)

Full-length cDNA libraries and normalization

REFERENCE

AUTHORS

Li,W.B., Gruber,C., Jesse,J. and Polayes,D.

TITLE

JOURNAL COMMENT

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Paraday Avenue Genoscope sequence ID : XCL0BA001ZE02PPL.

FEATURES

Location/Qualifiers

1..699

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="XCL0BA001ZE02"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT.6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

ORIGIN

Query Match 4.9%; Score 42.4; DB 13; Length 699;

Best Local Similarity 16.4%; Pred. No. 22;

Matches 44; Conservative 113; Mismatches 112; Indels 0; Gaps 0;

QY 95 TCGGGCGGCTCGGGCGCAATACGGGACACCGAGCGTGGATTTTCGCCGCTGGCGCGT 154

DB 670 BCCGGGGGGGGCCCCCCKKGGGKKKGGGGBBGGGGGGGGGKKKKKKKKKKKKKKKK 611

QY 155 TCGGGCGCAAGCTGATCTGTTCCCGCTGTTGGGTTTCGGCGCAGCAGCATTTGTCAGCC 214

DB 610 CCKKKKKGGGKKKCCCCCCCCCKKKKKKKKKKKGGGGGGGGGGGGGGGGGGGGGGGG 551

QY 215 CGCTGTTCACGCCCAAGGTGTGGCGCTGGATCAAGCTGTGTGGCAGCTTGATGACCG 274

DB 550 KKK 491

QY 275 CATTTGGCCATCAAACTGATGTTGATGGGTTAGTTTTCGGGGTTCGGAATCGTGGCGCT 334

DB 490 KKK 431

QY 335 TCGGCCAAATGTTGATGCGCGCGCTGCTGG 363

DB 430 CCGCCCCCKKKKKKKKKGGGKTTTGGG 402

RESULT 10

BY246241/c

LOCUS

DEFINITION

BY246241 RIKEN full-length enriched, visual cortex Mus musculus

CDNA clone K230304B15 5', mRNA sequence.

ACCESSION

BY246241

VERSION

BY246241.1 GI:26427753

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 457)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schönbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,

Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Mateuda, H.,

Batalov, S., Bessel, K. W., Blake, J. A., Bradt, D., Brusic, V.,

Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,

Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,

Gariboldi, M., Giesi, C., Godzik, J., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,

Kawaji, H., Kawasaki, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,

Kondo, S., Kondo, H., Kondo, S., Kondo, S., Kondo, S., Kondo, S.,

Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S.,

Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S.,

Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S.,

Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S.,

Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S., Kondo, S.,

Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., Mckenzie, D., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Read, J. C., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, K., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354583

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Itoh, M., Kawai, J., Konno, H., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Michela Pagliolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiroawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

200 CCTGTGCTCGCCACAGAAAGCTCAGTCTGGTGGAGGACAGTGGCTGGTATTCGGAC 141

440 GGTCACTGACGAGACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 499

140 CCTCGGTGTGGGCGCCAGAGTCCAGGCAATCTCTTCTCAGCCGCCCGCTTGGGTT 81

500 GCTGCGCAGCACCACCAAGCCCATCTGGGCAAGTGAATGCTGCCCGCTTCTCTGGGCGA 559

80 CATGGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 21

560 TGTGAT 565

20 CGGCT 15

CD861975 538 bp mRNA linear EST 11-JUL-2003

AZ01.101N14F010125 AZ01 Triticum aestivum cDNA clone AZ01101N14, mRNA sequence.

CD861975

CD861975.1 GI:32545791

EST.

Triticum aestivum (bread wheat)

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae; Triticeae; Triticum.

1 (bases 1 to 538)

Genoplatte.

Genoplatte, a major partnership french program in plant genomics

Unpublished (2003)

Contact: Genoplatte

Genoplatte

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (http://www.genoplatte.com and http://genoplatte-info.infobiogen.fr).

FEATURES

source

1. 538

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="recital"

/db_xref="taxon:4565"

/clone="AZ01101N14"

/tissue_type="leaf"

/clone_lib="AZ01"

Query Match 4.8%; Score 41.8; DB 14; Length 538;

Best Local Similarity 47.2%; Pred. No. 27;

Matches 127; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

1 GTGCAATAGCAGCGGGTTGGTAAAGCCCATGTTGATGCAATCGTGTGCTGACCTGGTTG 60

109 GTGAGGAGAAAGCTGTGGCGCTCAAGCCCAAGAGCATCGACTTCGGCGCAGCGCCGGT 168

61 AACCCGATGATTTTGGACGGCTTGTGTTTATCGCGCGGTGGCGGCAATACGGC 120

169 CTGCGCGCTGCGCAATCGAGAGACCCCGCATGAGGCGCTCGAAGAGAGAGGCTTCTCCGCGCGC 228

121 GACACCGGACGGTGGATTTTCGCCCGCTTGGCGCGCTTGGCGGCAAGCTGATCTGTTGGTCCG 180

229 AAGTCCATCTTGTCTCGCGCGCGCGCGGAGTTCGGGACCTTGTTCATCCAGCTGGCA 288

181 CTGTGGGGTTTGGCGCAGCAGCATTTGTCAGCCCGCTGTCAGCCCGGAGGTGGCGC 240

289 AAGCAAGTTTACGGCGCATCGAAAGTGGCGGCGCACAGCCAGCAGCCCCAAGAGCTGGAGTC 348

241 TGGATCAAGCTGCTGTGGGAGCTTGTGAT 269

Query Match 4.8%; Score 42; DB 13; Length 457;

Best Local Similarity 51.6%; Pred. No. 22;

Matches 96; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

380 CTCCAACTCGCGCGTCAGAAACTCAAGTTGTGATGATCAAGGCTGTGTGACGT 439

FEATURES

source

1. 457

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="K230304B15"

/tissue_type="visual cortex"

/clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

sequence that is located between the Not I site and the cgg18 tail. The sequence tag for this library is CCGGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

Db	477	CCACCGTGTGACATGCGACACCGTCACGTTGATGCCGTTCTCCAGCGCTTGTCCGCGCGC	418
Qy	218	TGTCAGCCCAAGGTGTGGCGCTGGATCAACGTGCTCGTGCAGTGTGTGATGACCGCAT	277
Db	417	TGCGTCAACCGCAGCGCGCCACGAGTGCCTCTCGGCGCGTGTGTCTTGAATGAGTAGTGGC	358

Qy	278	TGGCCATCAAACTGATGATTCATAGGTGTAGTTTTCGGGCGTTTGGAAATCGGTGGCCTTCG	337
Db	357	TGTCTTGGCCCTCGAATGGTGAATGCAGGTTCTCAGGTAGACGGCGCTTGTGTGACGACGG	298
Qy	338	CCCAAAATGTTGATGCCGCCCTCGTGGAAATCT	370
Db	297	CGGCACTTTCATGTCAGTCTCTCGTTGGCGATGT	265

DEFINITION	601681396F1_NTH_MGC_9 Homo sapiens cDNA clone IMAGE:3951622 5', mRNA sequence.
ACCESSION	BE998495
VERSION	BE998495.1 GI:10365029
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLNL at: image.lnl1.gov
plate: LICM819 row: p column: 23
High quality sequence stop: 765.
Location/Qualifiers
1. .770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3951622"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 9"
/notes="Organ: ovary; Vector: pOTB7; site 1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:03:00 ; Search time 60 Seconds
(without alignments)

1111.354 Million cell updates/sec

Title: US-09-105-117K-2

Perfect score: 1191

Sequence: 1 MVMEIFITGLLGASLLS.....INVVAVVMTALAIXIMMG 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1191	100.0	236	2	AAW37715 C. glutam
2	1191	100.0	236	4	AAW37715 C. glutam
3	1191	100.0	236	6	ABR58213 LysE prot
4	1178	98.9	233	4	AAW79660 Coryneb
5	1178	98.9	233	4	AAU71888 C. glutam
6	1178	98.9	233	7	ADB66197 Protein e
7	864.5	72.6	231	4	AAW64047 Coryneb
8	584	49.0	124	6	ABR58214 LysE24 pr
9	350	29.4	211	3	AAW01789 Escherich
10	298	25.0	208	6	ABM68839 Phototrab
11	298	25.0	211	6	ADA34450 Acinetoba
12	170.5	14.3	210	2	AAW20426 H. pylori
13	170.5	14.3	215	2	AAW20968 H. pylori
14	164	13.8	205	6	ABM72206 Staphyloc
15	132.5	11.1	205	6	ABM67507 Phototrab
16	132.5	11.1	226	4	AAW9911 C. glutam
17	121.5	10.2	235	6	ADA36962 Acinetoba
18	119.5	10.0	173	5	ABP39181 Staphyloc
19	119.5	10.0	222	6	ADA35184 Acinetoba
20	119.5	10.0	620	6	ABU49694 Protein e
21	114.5	9.6	229	3	ABU1637 A. vitis
22	112	9.4	203	6	ABM70439 Phototrab
23	108.5	9.1	222	6	ABM68944 Phototrab
24	108.5	9.1	238	6	ADA35835 Acinetoba
25	105	8.8	206	3	AAW99598 E. coli L

26	103.5	8.7	181	6	ABU16949
27	101.5	8.5	220	6	ADA33616 Acinetoba
28	101	8.5	201	6	ABM68999 Phototrab
29	101	8.5	212	3	AAW01787 Escherich
30	99	8.3	208	6	ADA33824 Acinetoba
31	96	8.1	213	6	ABM67782 Phototrab
32	95	8.0	295	6	ABU49858 Protein e
33	95	8.0	390	4	AAU58229 Propionib
34	95	8.0	390	6	ABM54748 Propionib
35	94.5	7.9	240	6	ADA36250 Acinetoba
36	93.5	7.9	452	4	AAU34590 E. coli c
37	93.5	7.9	452	6	ABU28648 Protein e
38	93	7.8	211	6	ABU19902 Protein e
39	93	7.8	459	4	AAW78987 C. glutam
40	93	7.8	459	4	AAW93244 C. glutam
41	92	7.7	223	3	AAW01786 Escherich
42	92	7.7	694	5	ABW93711 Herbicida
43	91.5	7.7	446	4	AAW76722 Coryneb
44	91.5	7.7	446	4	AAW92602 C. glutam
45	91.5	7.7	446	5	AAW22703 Coryneb

ALIGNMENTS

RESULT 1
AAW37715
ID AAW37715 standard; protein; 236 AA.
XX
AC AAW37715;
XX
DT 12-MAR-1998 (first entry)
XX
DE C. glutamicum Lys E protein (lysine export protein).
XX
LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;
XX
Microbial production; amino acid; animal feed additive.
XX
Corynebacterium glutamicum.
XX
DE19548222-A1.
XX
26-JUN-1997.
XX
22-DEC-1995; 95DE-01048222.
XX
22-DEC-1995; 95DE-01048222.
XX
(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
Vrijic M, Eggeling L, Sahm H;
XX
WPI; 1997-333867/31.
XX
N-PSDB; AAU96816.
XX
Increasing microbial production of amino acids, especially lysine - by
PT improving export carrier activity or corresponding gene expression, also
PT new export and regulatory genes from Corynebacterium.
XX
Claim 42; Page 10; 16pp; German.
XX
This sequence is the LysE protein product, a lysine export protein. LysG
and LysE encode a lysine transport regulatory protein and an export
protein, respectively. Microbial production of amino acids (A) is
CC improved by increasing the export-carrier activity and/or the export gene
CC expression in a microorganism that produces (A). The method is
CC specifically used to increase production of lysine, used as an animal
CC feed additive. Other (A) are variously useful as pharmaceuticals,
CC condiments and intermediates for fine chemicals. This method increases
CC the amount of (A) secreted into the culture medium. Export of (A) has
XX been found to depend on a single gene
SQ Sequence 236 AA;

Query Match 100.0%; Score 1191; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 4.6e-128; Indels 0; Gaps 0;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWMEIFITGLLGGASLLSIGPQNVVVKQIKREGLIAVLLVCLISDVFLFIAGTLGV 60
DB 1 MWMEIFITGLLGGASLLSIGPQNVVVKQIKREGLIAVLLVCLISDVFLFIAGTLGV 60
QY 61 DLLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVEAPQIIEETEPTVPDDTPLGGS 120
DB 61 DLLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVEAPQIIEETEPTVPDDTPLGGS 120
QY 121 AVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDAFVFIGGVAQYGDGTGRW 180
DB 121 AVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDAFVFIGGVAQYGDGTGRW 180
QY 181 IFAGAFAASLIWPLVGFGAALSRPLSSPKVWRINVVAVVMTALAIKMLMG 236
DB 181 IFAGAFAASLIWPLVGFGAALSRPLSSPKVWRINVVAVVMTALAIKMLMG 236

RESULT 2
AAG93201
ID AAG93201 standard; protein; 236 AA.
XX AC AAG93201;
XX DT 26-SEP-2001 (first entry)
XX DE C glutamicum protein fragment SEQ ID NO: 6955.
XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX KM organic acid synthesis.
XX OS Corynebacterium glutamicum.
XX PN EP1108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-00127688.
XX PR 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-00159162.
XX PR 03-AUG-2000; 2000JP-00280988.
XX PA (KYOW) KYOWA HAKKO KOGYO KK.
XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX DR WPI: 2001-376931/40.
XX DR N-PSDB; AAH68420.

Novel polynucleotides derived from Corynebacterium bacteria, for identifying mutation point of a gene, measuring expression of a gene, analyzing expression profile or pattern of a gene and identifying homologous gene.
PS Claim 29; SEQ ID NO 6955; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein sequences from the Corynebacterium bacteria Corynebacterium glutamicum. These are useful for identifying the mutation point of a gene derived from a mutant of corynebacterium bacteria, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from Corynebacterium bacteria, and identifying a homologue of a gene derived from Corynebacterium bacteria. Corynebacterium bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a protein described in the exemplification of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office

XX SQ Sequence 236 AA;
Query Match 100.0%; Score 1191; DB 4; Length 236;
Best Local Similarity 100.0%; Pred. No. 4.6e-128; Indels 0; Gaps 0;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWMEIFITGLLGGASLLSIGPQNVVVKQIKREGLIAVLLVCLISDVFLFIAGTLGV 60
DB 1 MWMEIFITGLLGGASLLSIGPQNVVVKQIKREGLIAVLLVCLISDVFLFIAGTLGV 60
QY 61 DLLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVEAPQIIEETEPTVPDDTPLGGS 120
DB 61 DLLSNAAPVLDIMRWGGIAYLLWFAVMAAKDAMTKVEAPQIIEETEPTVPDDTPLGGS 120
QY 121 AVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDAFVFIGGVAQYGDGTGRW 180
DB 121 AVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLPNAYLDAFVFIGGVAQYGDGTGRW 180
QY 181 IFAGAFAASLIWPLVGFGAALSRPLSSPKVWRINVVAVVMTALAIKMLMG 236
DB 181 IFAGAFAASLIWPLVGFGAALSRPLSSPKVWRINVVAVVMTALAIKMLMG 236

RESULT 3
ABR58213
ID ABR58213 standard; protein; 236 AA.
XX AC ABR58213;
XX DT 23-OCT-2003 (revised)
XX DT 11-AUG-2003 (first entry)
XX DE LysE protein.
XX KW L-lysine; L-arginine, LysE.
XX OS Corynebacterium glutamicum.
XX PN EP1266966-A2.
XX PD 18-DEC-2002.
XX PF 05-JUN-2002; 2002EP-00012539.
XX PR 12-JUN-2001; 2001JP-00177075.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Gunji Y, Yasueda H;
XX DR WPI: 2003-241171/24.
XX DR N-PSDB; ACC80941.
XX PT Novel DNA encoding variant of LysE protein from a coryneform bacterium, when introduced into methanol assimilating bacterium, facilitates excretion of L-lysine and/or L-arginine to outside of a cell.
XX PS Disclosure; Page 18-19; 23pp; English.
XX The present invention relates to DNA encoding variants of protein with loop region and six hydrophobic helices which facilitates excretion of L-lysine and/or L-arginine to outside of cell of a methanol assimilating bacterium when introduced into the bacterium. The method is used for encoding a protein which facilitates excretion of L-lysine, L-arginine or both of these L-amino acids to outside of a cell of a methanol assimilating bacterium when DNA of the method is introduced into the bacterium. The present sequence represents a lysE protein from Brevibacterium lactofermentum. (Updated on 23-OCT-2003 to standardise OS field)
XX SQ Sequence 236 AA;

Query Match 100.0%; Score 1191; DB 6; Length 236;
 Best Local Similarity 100.0%; Pred. No. 4.6e-128;
 Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYWEIFITGLLGASLLSISGPQNVLVKQIKREGIAVLVCLISDFVFIAGTGLV 60
 DB 1 MYWEIFITGLLGASLLSISGPQNVLVKQIKREGIAVLVCLISDFVFIAGTGLV 60

QY 61 DILSNAAPVLDIMRWGGIAYLLWFAVMAAKDANTNKVEAPQIIETEPTVDDTPLGGS 120
 DB 61 DILSNAAPVLDIMRWGGIAYLLWFAVMAAKDANTNKVEAPQIIETEPTVDDTPLGGS 120

QY 121 AVATDTRNRVRVEVSDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRW 180
 DB 121 AVATDTRNRVRVEVSDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRW 180

QY 181 IPAAGAFASLIWFFLVGFGAALSRPLSPKVRWVNVVAVMTALAKIMLMG 236
 DB 181 IPAAGAFASLIWFFLVGFGAALSRPLSPKVRWVNVVAVMTALAKIMLMG 236

RESULT 4
 AAB79660
 ID AAB79660 standard; protein; 233 AA.
 AC AAB79660;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:54.
 XX
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
 KW fine chemical production; microorganism; organic acid; nucleoside;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN WC200100843-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-IB000923.
 XX
 PR 25-JUN-1999; 99US-0141031P.
 PR 01-JUL-1999; 99DE-01030476.
 PR 02-JUL-1999; 99US-0142101P.
 PR 08-JUL-1999; 99DE-01031415.
 PR 08-JUL-1999; 99DE-01031418.
 PR 08-JUL-1999; 99DE-01031419.
 PR 08-JUL-1999; 99DE-01031420.
 PR 08-JUL-1999; 99DE-01031424.
 PR 08-JUL-1999; 99DE-01031428.
 PR 08-JUL-1999; 99DE-01031434.
 PR 08-JUL-1999; 99DE-01031435.
 PR 08-JUL-1999; 99DE-01031443.
 PR 08-JUL-1999; 99DE-01031453.
 PR 08-JUL-1999; 99DE-01031457.
 PR 08-JUL-1999; 99DE-01031465.
 PR 08-JUL-1999; 99DE-01031478.
 PR 08-JUL-1999; 99DE-01031510.
 PR 08-JUL-1999; 99DE-01031541.
 PR 08-JUL-1999; 99DE-01031573.
 PR 08-JUL-1999; 99DE-01031592.
 PR 08-JUL-1999; 99DE-01031634.
 PR 08-JUL-1999; 99DE-01031636.
 PR 08-JUL-1999; 99DE-01031636.
 PR 09-JUL-1999; 99DE-01032125.
 PR 09-JUL-1999; 99DE-01032126.
 PR 09-JUL-1999; 99DE-01032130.
 PR 09-JUL-1999; 99DE-01032186.
 PR 09-JUL-1999; 99DE-01032206.

PR 09-JUL-1999; 99DE-01032227.
 PR 09-JUL-1999; 99DE-01032228.
 PR 09-JUL-1999; 99DE-01032229.
 PR 09-JUL-1999; 99DE-01032230.
 PR 14-JUL-1999; 99DE-01032922.
 PR 14-JUL-1999; 99DE-01032926.
 PR 14-JUL-1999; 99DE-01032928.
 PR 14-JUL-1999; 99DE-01033004.
 PR 14-JUL-1999; 99DE-01033005.
 PR 14-JUL-1999; 99DE-01033006.
 PR 12-AUG-1999; 99US-0148613P.
 PR 27-AUG-1999; 99DE-01040764.
 PR 27-AUG-1999; 99DE-01040765.
 PR 27-AUG-1999; 99DE-01040766.
 PR 27-AUG-1999; 99DE-01040832.
 PR 31-AUG-1999; 99DE-01041378.
 PR 31-AUG-1999; 99DE-01041379.
 PR 31-AUG-1999; 99DE-01041380.
 PR 31-AUG-1999; 99DE-01041394.
 PR 03-SEP-1999; 99DE-01042076.
 PR 03-SEP-1999; 99DE-01042077.
 PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042086.
 PR 03-SEP-1999; 99DE-01042087.
 PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042124.
 PR 03-SEP-1999; 99DE-01042129.
 PR 09-MAR-2000; 2000US-0187970P.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompeius M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 XX WPI; 2001-137957/14.
 DR N-ESDB; AAF71779.
 XX
 PT Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway
 PT proteins, useful for producing fine chemicals in microorganisms,
 PT including organic acids, nonproteinogenic amino acids, and purine and
 PT pyrimidine bases.
 XX
 PS Claim 20; Page 234-235; 1737pp; English.
 XX
 CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
 CC nucleic acids are useful for the production of fine chemicals in
 CC microorganisms, including organic acids, nonproteinogenic amino acids,
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
 CC vitamins, cofactors, polyketides and enzymes
 XX
 SQ Sequence 233 AA;

Query Match 98.9%; Score 1178; DB 4; Length 233;
 Best Local Similarity 100.0%; Pred. No. 1.4e-126;
 Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEIFITGLLGASLLSISGPQNVLVKQIKREGIAVLVCLISDFVFIAGTGLV 63
 DB 1 MEIFITGLLGASLLSISGPQNVLVKQIKREGIAVLVCLISDFVFIAGTGLV 60

QY 64 SNAAPVLDIMRWGGIAYLLWFAVMAAKDANTNKVEAPQIIETEPTVDDTPLGSSAVA 123
 DB 61 SNAAPVLDIMRWGGIAYLLWFAVMAAKDANTNKVEAPQIIETEPTVDDTPLGSSAVA 120

QY 124 TDTNRNRVRVEVSDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRWIFA 183
 DB 121 TDTNRNRVRVEVSDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRWIFA 180

QY 184 AGAPAFASLIWFFLVGFGAALSRPLSPKVRWVNVVAVMTALAKIMLMG 236

Db 181 AGAFAASLIWFLVGFAGAAALSRPLSPKVRWVWVAVVMTALAIKMLMG 233

RESULT 5
AAU71888
AC AAU71888 standard; protein; 233 AA.
AC AAU71888;
DT 26-FEB-2002 (first entry)
DE C. glutamicum metabolic pathway protein encoded by gene #23.
XX
XX Metabolic pathway protein; MP; lysine biosynthesis pathway;
KW methionine biosynthesis pathway; large-scale production of fine chemical;
KW Corynebacterium diphtheriae; diphtheria.
XX
XX Corynebacterium glutamicum.
OS
XX
XX WO2001:66573-A2.
PN
XX
XX 13-SEP-2001.
PD
XX
XX 22-DEC-2000; 2000WO-IB002035.
PF
XX
XX 09-MAR-2000; 2000US-0187970P.
PR
XX
XX 23-JUN-2000; 2000US-00606740.
XX
XX (BADI) BASF AG.
PA
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
PI Kim J, Lee H, Hwang B;
PI
XX
XX WPI; 2001-582369/65.
DR N-PSDB; AAS96098.
DR
XX
XX Nucleic acids encoding metabolic pathway proteins from Corynebacterium
PT glutamicum, useful for producing methionine and lysine in Corynebacterium
PT and Brevibacterium.
PT
XX
XX Disclosure; Page 215-216; 316pp; English.
PS
XX
XX The present invention relates to the isolation of novel Corynebacterium
CC glutamicum genes (AAS96073-AAS96132) encoding metabolic pathway (MP)
CC proteins. The metabolic pathway proteins of the invention include enzymes
CC involved in the lysine and methionine biosynthetic pathways. The
CC polynucleotide sequences of the invention can be used for the large-scale
CC production and/or modulation of expression of fine chemicals such as
CC lysine and methionine. The sequences of the invention may be used to
CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a
CC subject to detect diphtheria. AAU71863-AAU71922 represent the novel C.
CC glutamicum metabolic pathway proteins of the invention
XX
SQ Sequence 233 AA;
Query Match 98.9%; Score 1178; DB 4; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.4e-126;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 MEIITGILLGASLLSISGPNVLIKQIKREGILAVLVCLISDVFLFIAGTGLVDLL 63
Db 1 MEIITGILLGASLLSISGPNVLIKQIKREGILAVLVCLISDVFLFIAGTGLVDLL 60
QY 64 SNAAPIVLDIMRWGGIAYLLWFVAVMAAKDAMTNKVEAPQIIEETPTVPDDTPLGGSAVA 123
Db 61 SNAAPIVLDIMRWGGIAYLLWFVAVMAAKDAMTNKVEAPQIIEETPTVPDDTPLGGSAVA 120
QY 124 TDTNRNRVVEVSDQRVWKPMMLAIVLTWLNPNAYLDAFVFIGVGAQYGDGTGRWIFA 183
QY 121 TDTNRNRVVEVSDQRVWKPMMLAIVLTWLNPNAYLDAFVFIGVGAQYGDGTGRWIFA 180
QY 184 AGAFAASLIWFLVGFAGAAALSRPLSPKVRWVWVAVVMTALAIKMLMG 236
Db 181 AGAFAASLIWFLVGFAGAAALSRPLSPKVRWVWVAVVMTALAIKMLMG 233

Db 181 AGAFAASLIWFLVGFAGAAALSRPLSPKVRWVWVAVVMTALAIKMLMG 233

RESULT 6
ADB66197
ID ADB66197 standard; protein; 233 AA.
XX
AC ADB66197;
XX
DT 04-DEC-2003 (first entry)
DE Protein encoded by C. glutamicum LysE gene.
XX
XX L-arginine production; coryneform bacteria; lysE; arginine repressor;
KW argR; liver function promoting agent; amino acid infusion;
KW amino acid pharmaceutical.
XX
XX Corynebacterium glutamicum.
OS
XX
XX US2003113899-A1.
PN
XX
XX 19-JUN-2003.
PD
XX
XX 17-JUL-2002; 2002US-00196232.
PF
XX
XX 25-JUL-2001; 2001JP-00224586.
PR
XX
XX (AJIN) AJINOMOTO CO INC.
PA
XX
XX Yamaguchi M, Ito H, Gunji Y, Yasueda H;
PI
XX
XX WPI; 2003-708853/67.
DR N-PSDB; ADB66196.
DR
XX
XX A microorganism comprising enhanced expression of the lysE gene is useful
PT for enhanced production of L-arginine.
PT
XX
XX Disclosure; Page 27-28; 36pp; English.
PS
XX
XX The present invention relates to a method for producing L-arginine in a
CC microorganism (e.g. coryneform bacteria) that has L-arginine producing
CC ability and has been modified for enhanced expression of the lysE gene.
CC The microorganism is also modified so that an arginine repressor (argR)
CC does not function normally. The method of the invention is useful for the
CC enhanced production of L-arginine which is useful in liver function
CC promoting agents, amino acid infusion and comprehensive amino acid
CC pharmaceuticals. The present sequence represents a protein encoded by the
CC Corynebacterium glutamicum LysE gene. Note: The present sequence is given
CC as SEQ ID No:25 in the Sequence Listing but is referred to as SEQ ID
CC No:26 in the rest of the specification.
XX
SQ Sequence 233 AA;
Query Match 98.9%; Score 1178; DB 7; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.4e-126;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 MEIITGILLGASLLSISGPNVLIKQIKREGILAVLVCLISDVFLFIAGTGLVDLL 63
Db 1 MEIITGILLGASLLSISGPNVLIKQIKREGILAVLVCLISDVFLFIAGTGLVDLL 60
QY 64 SNAAPIVLDIMRWGGIAYLLWFVAVMAAKDAMTNKVEAPQIIEETPTVPDDTPLGGSAVA 123
Db 61 SNAAPIVLDIMRWGGIAYLLWFVAVMAAKDAMTNKVEAPQIIEETPTVPDDTPLGGSAVA 120
QY 124 TDTNRNRVVEVSDQRVWKPMMLAIVLTWLNPNAYLDAFVFIGVGAQYGDGTGRWIFA 183
QY 121 TDTNRNRVVEVSDQRVWKPMMLAIVLTWLNPNAYLDAFVFIGVGAQYGDGTGRWIFA 180
QY 184 AGAFAASLIWFLVGFAGAAALSRPLSPKVRWVWVAVVMTALAIKMLMG 236
Db 181 AGAFAASLIWFLVGFAGAAALSRPLSPKVRWVWVAVVMTALAIKMLMG 233

RESULT 7

AA64047
ID AAG64047 standard; protein; 231 AA.
XX
AC AAG64047;
XX
DT 11-SEP-2001 (first entry)
XX
DE Corynebacterium thermoaminogenes lysin biosynthetic enzyme lyse.
XX
KW Heat-resistant; lysin biosynthesis; enzyme; coryneform;
KW aspartate-semialdehyde dehydrogenase; lyse.
XX
OS Corynebacterium thermoaminogenes.
XX
PN JP2001120270-A.
XX
PD 08-MAY-2001.
XX
PF 01-NOV-1999; 99JP-00311148.
XX
PR 01-NOV-1999; 99JP-00311148.
XX
PA (AJIN) AJINOMOTO KK.
XX
DR WPI; 2001-364760/38.
DR N-PSDB; AAH45375.
XX

XX A heat-resistant lysin biosynthetic system enzyme gene of a high
XX temperature-resistant coryneform microbe.
XX
XX Claim 7; Page 24; 27pp; Japanese.
XX
XX The invention relates to a gene from a high temperature-resistant
XX coryneform microbe that encodes a heat-resistant lysin biosynthetic
XX enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and
XX can be used for growing amino acid-producing microbes. The present amino
XX acid sequence corresponds to an enzyme of the invention
XX
XX Sequence 231 AA;

Query Match 72.6%; Score 864.5; DB 4; Length 231;
Best Local Similarity 71.2%; Pred. No. 1.4e-90;
Matches 166; Conservative 30; Mismatches 32; Indels 5; Gaps 2;

Qy 4 MEIPIITGLLGASLLISIGPQNVIVIKQIKREGIIVLVCLISDVFLIAGTGLV 63
Db 4 MEIPIITGLLGASLLISIGPQNVIVIKQIKREGIIVLVCLISDVFLIAGTGLV 63
Qy 64 SNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETPTVDDTPLGSSAVA 123
Db 64 SDTAPIILDLRWCGIAYLLWFVMAAKDAMTNKVEAPQIIEETPTVDDTPLGSSAVA 118
Qy 124 TDTNRVVEVSVQKRVKPMMAIVLTWLNPNAYLDAPFVIGGGAQYGDTRWIFA 183
Db 119 VTTQRPRLITSGTRQVVRPMMAIVLTWLNPNAYLDAPFVIGGGAQYGDTRWIFA 178
Qy 184 AGAFAASLWFLVGFAGAAALSRPLSPRVWENINIVAVVMTALAKMLMG 236
Db 179 AGAFAASLWFLVGFAGAAALSRPLSPRVWENINIVAVVMTALAKMLMG 231

RESULT 8

ABR58214
ID ABR58214 standard; protein; 124 AA.
XX
AC ABR58214;
XX
DT 23-OCT-2003 (revised)
DT 11-AUG-2003 (first entry)
XX
DE Lyse24 protein.

XX L-lysine; L-arginine; Lyse24.
KW Corynebacterium glutamicum.
OS
PN EPI266966-A2.
XX
PD 18-DEC-2002.
XX
PF 05-JUN-2002; 2002EP-00012539.
XX
PR 12-JUN-2001; 2001JP-00177075.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Gunji Y, Yasueda H;
XX
DR WPI; 2003-241171/24.
DR N-PSDB; ACC80942.
XX

XX Novel DNA encoding variant of Lyse protein from a coryneform bacterium,
XX when introduced into methanol assimilating bacterium, facilitates
XX excretion of L-lysine and/or L-arginine to outside of a cell.

XX Claim 8; Page 20; 23pp; English.

XX The present invention relates to DNA encoding variants of protein with
XX loop region and six hydrophobic helices which facilitates excretion of L-
XX lysine and/or L-arginine to outside of cell of a methanol assimilating
XX bacterium when introduced into the bacterium. The method is used for
XX encoding a protein which facilitates excretion of L-lysine, L-arginine or
XX both of these L-amino acids to outside of a cell of a methanol
XX assimilating bacterium when DNA of the method is introduced into the
XX bacterium. The present sequence represents a Lyse24 protein from
XX Brevibacterium lactofermentum. (Updated on 23-OCT-2003 to standardise OS
XX field)

XX Sequence 124 AA;

Query Match 49.0%; Score 584; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 1e-58;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYMEIITGLLGASLLISIGPQNVIVIKQIKREGIIVLVCLISDVFLIAGTGLV 60
Db 1 MYMEIITGLLGASLLISIGPQNVIVIKQIKREGIIVLVCLISDVFLIAGTGLV 60
Qy 61 DLLSNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETPTVDDTPLG 118
Db 61 DLLSNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETPTVDDTPLG 118

RESULT 9

AA01789
ID AA01789 standard; protein; 211 AA.
XX
AC AA01789;
XX
DT 03-JAN-2001 (first entry)
XX
DE Escherichia coli YggA amino acid excretion protein.
XX
KW E. coli; yggA gene; amino acid production; excretion protein gene;
KW amino acid excretion protein.
XX
OS Escherichia coli.
XX
PN EPI016710-A2.
XX
PD 05-JUL-2000.
XX
PF 17-DEC-1999; 99EP-00125263.
XX

XX WO9640893-A1.
 XX 19-DEC-1996.
 XX 06-JUN-1996; 96WO-US009122.
 XX 07-JUN-1995; 95US-00487032.
 XX 01-APR-1996; 96US-00630405.
 XX (ASTR) ASTRA AB.
 XX Smith D, Berglindh OT, Mellgaerd BL;
 XX WPI: 1997-052306/05.
 XX N-PSDB; AAT68221.
 XX Helicobacter pylori nucleic acid sequences and related polypeptide(s) -
 PT useful for vaccines to treat or prevent H. pylori infection, and to
 PT detect Helicobacter.
 XX
 XX Claim 73; Page 1359; 1481pp; English.
 XX This sequence represents a H. pylori protein likely to contain five
 CC membrane spanning regions. The protein may be used in a vaccine to
 CC prevent or treat H. pylori infection or to identify H. pylori polypeptide
 CC binding compounds, useful as potential H. pylori life cycle activators or
 CC inhibitors. The genomic sequence of H. pylori (ATCC 55679) was determined
 CC from overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides, and
 CC the predicted coding regions defined by computer evaluation. To identify
 CC likely H. pylori antigens for vaccine development, the amino acid
 CC sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts
 XX
 XX Sequence 215 AA;

Query Match 14.3%; Score 170.5; DB 2; Length 215;
 Best Local Similarity 21.5%; Pred. No. 6.7e-11;
 Matches 50; Conservative 48; Mismatches 92; Indels 43; Gaps 5;
 QY 6 IFTGGLGASLLSIGPQNVLIKQIKREGIAVLVCLISDVFLFIAGTLGVLLSN 65
 Db 9 VFIEGGLAISLCAVAGQCSFIVERGMARNTVFILCALCMCDIVLMSMGVFGVGNYP- 67
 QY 66 AAPIVLDI-MRWGGIAYLLNFVAVMAAKDAMTKVEAPQIIETEPTVPDDTPLGGSVAT 124
 Db 68 AKNLYLSLFLNFAGVFTGFVFLALK-----T 95
 QY 125 DTRNRVREVSVDKORVWVKPMLMAIVLTWLNPNAYLDADFVFIGVGAGYQDGTGRWIFAA 184
 Db 96 LFQTFKKQVQTEPKLSLKKTLTFTLVTLNPQVLEWVFLIGASMSFNLVQKVFELA 155
 QY 185 GAFAASLWFPPL-----VGFGAALSRPLSSPKVWRINNVAVVMTALAIAKL 232
 Db 156 GTLSAAPSWMILLCTMSLRVG---SKLLNQKIFMGVNLVFTAINGTLSVTL 204

RESULT 14
 ABM72206
 ID ABM72206 standard; protein; 205 AA.
 XX
 XX ABM72206;
 XX 20-NOV-2003 (first entry)
 XX Staphylococcus aureus protein #1446.
 XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 KW enzymatic assay; antibiotic target.

XX Staphylococcus aureus.
 XX WO200294868-A2.
 XX 28-NOV-2002.
 XX 27-MAR-2002; 2002WO-IB002637.
 XX 27-MAR-2001; 2001GB-00007661.
 XX (CHIR-) CHIRON SPA.
 XX Masignani V, Mora M, Scarselli M;
 XX WPI: 2003-120786/11.
 XX N-PSDB; ACF73766.
 XX New Staphylococcus aureus protein, useful as a vaccine for treating or
 PT preventing Staphylococcal infection, specifically an infection caused by
 PT S. aureus, e.g. sepsis.
 XX
 XX Claim 1; SEQ ID NO 2892; 49pp; English.

XX The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus proteins of the invention

XX Sequence 205 AA;

Query Match 13.8%; Score 164; DB 6; Length 205;
 Best Local Similarity 21.4%; Pred. No. 3.5e-10;
 Matches 49; Conservative 47; Mismatches 93; Indels 40; Gaps 4;
 QY 8 ITGLLGASLLSIGPQNVLIKQI---KREGIAVLVCLISDVFLFIAGTLGVLLS 64
 Db 6 IHGFIILAIGLILFLGAQNVIFMQGANQPKYRYPALITAGLSLSLIIITAVGVSI 65
 QY 65 NAAPIVLDI-MRWGGIAYLLNFVAVMAAKDAMTKVEAPQIIETEPTVPDDTPLGGSVAT 124
 Db 66 MSLFVQLAIYIVGLIFLWYAW-----TIWEDK-----ST 97
 QY 125 DTRNRVREVSVDKORVWVKPMLMAIVLTWLNPNAYLDADFVFIGVGAGYQDGTGRWIFAA 184
 Db 98 DGETQIMSP-----MKQVSFALSLSLNPALITAGLSLSLIIITAVGVSI 148
 QY 185 GAFAASLWFPPLVGFGAALSRPLSSPKVWRINNVAVVMTALAIAKL 233
 Db 149 ACISVSWLWFLAILGKRWGSDIKTKLTTINKISSIIIIIVLMIL 197

RESULT 15
 ABM67507
 ID ABM67507 standard; protein; 205 AA.
 XX
 XX ABM67507;
 XX 20-NOV-2003 (first entry)
 XX Photorhabdus luminescens protein sequence #604.
 XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:24:51 ; Search time 22 Seconds
(without alignments)
553.806 Million cell updates/sec

Title: US-09-105-117K-2
Perfect score: 1191
Sequence: 1 MVIMEFITGILLGASLLLS.....INVVAVVMTALAIAKMLMG 236

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/2/iaa/5A COMB.pap.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pap.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pap.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pap.*
5: /cgn2_6/prodata/2/iaa/PCJUS COMB.pap.*
6: /cgn2_6/prodata/2/iaa/backfile1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	341	28.6	213	4	US-09-489-039A-11845
2	332	27.9	222	4	US-09-252-991A-24479
3	310	26.0	214	4	US-09-543-681A-4972
4	298	25.0	211	4	US-09-328-352-5737
5	138.5	11.6	212	4	US-09-543-681A-4767
6	135.5	11.4	260	4	US-09-252-991A-25992
7	121.5	10.2	235	4	US-09-328-352-8249
8	119.5	10.0	173	4	US-09-134-001C-4026
9	119.5	10.0	222	4	US-09-328-352-6271
10	118.5	9.9	250	4	US-09-543-681A-4487
11	117.5	9.7	206	4	US-09-252-991A-18746
12	115.5	9.7	219	4	US-09-489-039A-7731
13	113.5	9.5	228	4	US-09-543-681A-4854
14	111	9.3	222	4	US-09-252-991A-26644
15	108.5	9.1	238	4	US-09-328-352-7122
16	108	9.1	217	4	US-09-489-039A-8076
17	104.5	8.8	211	4	US-09-543-681A-6708
18	104	8.7	249	4	US-09-252-991A-25717
19	101.5	8.5	220	4	US-09-328-352-4903
20	99	8.3	208	4	US-09-328-352-5111
21	98.5	8.3	1107	4	US-09-489-039A-8890
22	96	8.1	1421	4	US-09-252-991A-17805
23	94.5	7.9	240	4	US-09-328-352-7537
24	94	7.9	211	4	US-09-489-039A-9608
25	94	7.9	552	4	US-09-489-039A-7419
26	93	7.8	228	4	US-09-252-991A-23837
27	91.5	7.7	484	4	US-09-634-238-216

28	91	7.6	211	4	US-09-252-991A-20747	Sequence 20747, A
29	90	7.6	205	4	US-09-489-039A-12661	Sequence 12661, A
30	89.5	7.5	205	4	US-09-396-357-2	Sequence 2, Appli
31	88.5	7.4	350	4	US-09-252-991A-32358	Sequence 32358, A
32	87	7.3	398	4	US-09-328-352-6411	Sequence 6411, Ap
33	87	7.3	451	4	US-09-107-532A-5352	Sequence 5352, Ap
34	86.5	7.3	226	4	US-09-252-991A-31610	Sequence 31610, A
35	86.5	7.3	456	4	US-09-543-681A-6928	Sequence 6928, Ap
36	86	7.2	485	4	US-09-134-001C-6031	Sequence 6031, Ap
37	85.5	7.2	224	4	US-09-134-001C-4608	Sequence 4608, Ap
38	85.5	7.2	269	4	US-09-489-039A-7320	Sequence 7320, Ap
39	85.5	7.2	472	4	US-09-252-991A-18544	Sequence 18544, A
40	85.5	7.2	475	4	US-09-252-991A-29697	Sequence 29697, A
41	84.5	7.1	384	4	US-09-252-991A-19333	Sequence 19333, A
42	84	7.1	214	4	US-09-328-352-6695	Sequence 6695, Ap
43	84	7.1	241	4	US-09-328-352-7159	Sequence 7159, Ap
44	84	7.1	289	4	US-09-328-352-6229	Sequence 6229, Ap
45	83.5	7.0	210	4	US-09-328-352-4775	Sequence 4775, Ap

ALIGNMENTS

RESULT 1
US-09-489-039A-11845
; Sequence 11845, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11845
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11845

Query Match 28.6%; Score 341; DB 4; Length 213;
Best Local Similarity 32.5%; Pred. No. 8.3e-32;
Matches 77; Conservative 43; Mismatches 83; Indels 34; Gaps 4;
QY 1 MVIMEFITGILLGASLLLSIGPQNVIVKQIKRGLIALLVCLISDVFLFIAGTIGV 60
DB 1 ILMFTYFQGLAQAAMILFLGQFNQAFVQIRQYHIMIALLCVSDLLICAGIFGG 60
QY 61 DLLSNAPIVLDIMRWGTAIVLLNFVAVMAAKDAMTKNVEAPQITEPTVDDTFLGGS 120
DB 61 SALLMQSPWLLAVTWGGVAFLLWYFGALKTAFPSQLEL-----100
QY 121 AVATDTRNREVEVSVDKQVVKPMLMAIVLTWLNPNAYLDADFVFGVGAQYG-DTGR 179
DB 101 -----ANAEVWQQRW-KIIITMLAVTWLNPHVYLDTFVLGSGGGLAVEPKR 148
QY 180 WIFAGAGAAASLIWFLVFGGAALSRLSPKPKWKNVNVVAVMTALAIAKMLMG 236
DB 149 W-FALGTISAFLLFFGLALLAALAPRLRTARAQRINIVVGAVMMFIAQLAREG 204

RESULT 2
US-09-252-991A-24479
; Sequence 24479, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

QY 122 VADTNRVRVEVDKQVWVKPMLMAIVLTWLNPNAYLDAFVIGV---GAQYGD 177
Db 121 -----TLLEALFQWINKSWMA---ISGITLYTSQYPIV 154
QY 178 GRWIFAGAFASLIWPLVFGAA---ALSRLSPKVRWVWVAVVMTALA KLM 233
Db 155 SMLLV---AIIFTLINPCVATWATGHSLEERKPKILKLFNFMGGLALSIVL 210

RESULT 6
US-09-252-991A-25992
; Sequence 25992, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25992
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25992

Query Match 11.4%; Score 135.5; DB 4; Length 260;
Best Local Similarity 23.7%; Pred. No. 1.3e-07;
Matches 57; Conservative 36; Mismatches 99; Indels 49; Gaps 9;

QY 4 MEIFTGILLGASLLISIGPQNVIVIKQI---KRGSLIAYL-LVCLISDVFLFAGTL- 58
Db 51 LSVLLTLAAVFAIALVSPGPDVALVVRTSLHQGRAGLASALGLACGI-----LLHTLV 105
QY 59 --GVOLLGNAAPVLDIMRWGGIAYLWFAVMAAKDAMTKVEAPQIIEETPTVDDTP 116
Db 106 LUGVSLLSRTPVLPAILQALGALYALWVGALRAWLRGGQGRGLDGLP----PSP 161
QY 117 LGGSAVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLNPNAYLDAFVIGV-GA QY 175
Db 162 LG-----PWLRGVATNLFNPKALVLFIALGSLIPAQMS 195
QY 176 DTGRWIFAGAFASLIWPLVFGAAALSRPLSSPKVR-----WVWVAVVMTALA KLM 232
Db 196 LGGKLAVALLFMGACWFGLL---SLTLTPALQARLLRAYPMLDAACGWFLVAAAI 252
QY 233 M 233
Db 253 L 253

RESULT 7
US-09-328-352-8249
; Sequence 8249, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8249
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8249

Query Match 10.2%; Score 121.5; DB 4; Length 235;
Best Local Similarity 23.0%; Pred. No. 4.8e-06;
Matches 57; Conservative 41; Mismatches 89; Indels 61; Gaps 12;

QY 3 IMEIFTGILLGASLLISIGPQNVIVIKQIKRGSLIAYLV---CLISDVFLFAGTLG 59
Db 29 LSQILAFGLIC-LAMVLTGPNMIVLISRSIS-QCKIAGFISIGGVAVGVFVYMLCASFG 86
QY 60 VDLISNAAPVLDIMRWGGIAYLWFAVMAAKDAMTKVEAPQIIEETPTVDDTP 119
Db 87 ITALWVAVPYAYDTIRIAGAWYLLMLAKALR-----PNAPIFN 126
QY 120 SAVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLNPN-----AYLDAFVP--IGVG 171
Db 127 VK-----DLAVDSP---LKLFLMGFLTNLANPKIAIMYLSLLPQFIHPQQSIL 172
QY 172 AQYGDTR-WIFAGAFASLIWPLVFGAAALSRPLSSPKVM-----RWNVVAVVMT 226
Db 173 AQSIQLGTIQIFVSVNA-----LIVFSAGSIALFLQKKPLMASIQRW---VMGTILA 223
QY 227 ALAIKML 234
Db 224 GLAVRIIL 231

RESULT 8
US-09-134-001C-4026
; Sequence 4026, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: CTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4026
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4026

Query Match 10.0%; Score 119.5; DB 4; Length 173;
Best Local Similarity 23.0%; Pred. No. 5.3e-06;
Matches 46; Conservative 32; Mismatches 83; Indels 39; Gaps 4;

QY 39 IAVLLVCLISDVFLFAGTLGVDLLSNAAPVLDIMRWGGIAYLWFAVMAAKDAMTKV 98
Db 8 LPVITTAGCUTFLIVITAILGVSLLIISMTPIQLFIYIGFLFMYM-----AMSLWT 60
QY 99 EAPQIIEETPTVDDTPLOGGSAVATDTRNRVRVEVDKQVWVKPMLMAIVLTWLNPN 158
Db 61 EKPSNIEETEP-----MSAKKQ-----ILFALSVSLLNPH 90
QY 159 AYLDAPVFIGVGAQYGDTRWIFAGAFASLIWPLVFGAAALSRPLSSPKVRWV 218
Db 91 AIMDTVGVTGSASVYDGVKVVFLSATISVSWTFVFLAILGRITGDKSGKYIILN 150
QY 219 VV--VAVVMTALA KLM 236
Db 151 KVSSVIVIIVGLIILKNIVG 170

RESULT 9
US-09-328-352-6471
; Sequence 6471, Application US/09328352
; Patent No. 6562958

```
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-039A
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6471
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-6471

Query Match          10.8%; Score 119.5; DB 4; Length 222;
Best Local Similarity 23.1%; Pred. No. 7.7e-06;
Matches 54; Conservative 38; Mismatches 95; Indels 47; Gaps 9;

QY 6 IFITGLLGASLLSIGPQNVLVKQIKREGLIALLVCLISDVFLFIAGTIGVDLNS 65
Db 24 IFIAFWSVILFIITPGADWAYAISAGIKGVVVPVAVAGMLFGHEFITILLVAGVGLVA 83
QY 66 AAPIVLDIMRGGIAYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPLGSSAVATD 125
Db 84 NNPTALMILTAVGSAIYLLMGI-----NLLTTP-----PT-PNKS-----GS----- 119
QY 126 TRNRVRVSVDKORVWVKPMLMAIIVLTWLPNAYLDAFVIGGVGAQYGDGTGRW----- 180
Db 120 -----EKAQSLRWATKGVVVSGLNPKVFL---LFLALLPQIDITASMSVTTQ 165
QY 181 IFAGA--FAASLIWFPLVGFGAAL--SRPLSPKVKRWVNVVAVMTALAI 230
Db 166 ILAFGVVHIISCAIYLMVGYGSEALLTRPQAALVGRFSGGLMVVATCLLI 219

RESULT 10
US-09-543-681A-4487
; Sequence 4487, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4487
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-4487

Query Match          9.9%; Score 118.5; DB 4; Length 250;
Best Local Similarity 21.9%; Pred. No. 1.2e-05;
Matches 52; Conservative 35; Mismatches 101; Indels 49; Gaps 8;

QY 6 IFITGLLGASLLSIGPQNVLVKQIKREGLIALLVCLISDVFLFIAGTIGVD 61
Db 53 IFIT-----LVPGENSIFVETSSAKGVK--GGYKAALGVTTGDALLIFIAFLGVA 101
QY 62 LLSNAAPVLDIMRGGIAYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPLGSSA 121
Db 102 SLVNTSPVFVIKYGALYLYTLGLKLYATFHKKKEQPE----- 142
QY 122 VATDTRNRVRVSVDKORVWVKPMLMAIIVLTWLPNAYL-DAFVIGGVGAQYGDGTGRW 180
Db 143 -----QVAVVWKAKNGLYIK-----ALFLSLNPKRMIIFFVSFFIQFIDPKYENAGVP 191
QY 181 IFAGAF--AASLIWFPLVGFGAALSRPLS--SPKVRWNVVAVVMTALAIKMLI 234
```

```
Db 192 PFVLGVILEICSMYLSVLIFGGVAITNKKRKLASLSNSCIGAVFLFGAKIAL 248

RESULT 11
US-09-252-991A-18746
; Sequence 18746, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18746
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-18746

Query Match          9.9%; Score 117.5; DB 4; Length 206;
Best Local Similarity 25.4%; Pred. No. 1.2e-05;
Matches 64; Conservative 31; Mismatches 82; Indels 75; Gaps 14;

QY 4 MEIFITGLLGASLLSIGPQNVLVKQIKREGLI-----VLLVCLISDVFLFIAGTIG 59
Db 10 MQSLVFPFLFAVVASITPGPTNLTIVLSNS-QRHGLAAAMPVILGCAVAALILLIG-LG 67
QY 60 VDLISNAAPVLDIMRW---GGIAYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTP 116
Db 68 LGELLRHPHPLQQGLAMLGWGLSYLANSLSRSAGG-----IDGAE-----P 110
QY 117 -----LGGSAVATDTRNRVRVSVDKORVWVKPMLMAIIVLTWLPNAY---LDAFVIG 168
Db 111 RRLGVLGAA-----LQLVNPKAMMALAALFA 140
QY 169 GVGQAQYDGTGR-WIFAGAFASLI-----IWPLVGFGAALSRPLSPKVRWNVVAV 223
Db 141 GEGA--GQAGRIGLLALLFFLNSLPCIASW-ALLGVGSARL-----LRSPSLMKRFNQGMAL 194
QY 224 VMTALAIKMLM 235
Db 195 LLASAWAALL 206

RESULT 12
US-09-489-039A-7731
; Sequence 7731, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7731
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-7731

Query Match          9.7%; Score 115.5; DB 4; Length 219;
Best Local Similarity 24.4%; Pred. No. 2.2e-05;
Matches 50; Conservative 37; Mismatches 99; Indels 19; Gaps 8;
```

```

RESULT 14
US-09-252-991A-26644
; Sequence 26644, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```

Qy	22	GPQN--VLVIKGGIKREGLIANVLLCLISDVFLPIAGTIGVDLISNAAPILVDIMRWGGI	79
Db	56	GPTNFIILSLSSHYKISKTLPTVLIGSCIGAAALLVLVVGIGLSTILAYPVIOKIMAWSG-	114
Qy	80	AVLLWFAVMAAKDAMTKNVEARPOIIEETPTVPDDPLGGSAVATDTRNVRVRESVDKQ	139
Db	115	--LWLTVLAWLKYNN--PVISLEKNQYPP--PIGFKA-----	147
Qy	140	RWVKPMLMAIVLTLNPNAYLDAFVFIGGYGAQYGDTCR-----WIFAAGAFAA	189
Db	148	-----AFLMQALNPKTMMAFAVI-----SVYTKQGDILNVVLSILCFLLIAEPC	194
Qy	190	SLIWFPIVWGGAALSRPLSGPKVWRWINVVAVVMTA	227
Db	195	LYLW-ALVG---RLSTRLLSKPKHINIFNKIMAILLLA	228

Search completed: March 12, 2004. 18:27:23

Thu Mar 18 12:34:41 2004

us-09-105-117k-2.ra1

Page 6

Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 06:36:47 ; Search time 43 Seconds
(without alignments)
1421.243 Million cell updates/sec

Title: US-09-105-117K-2

Perfect score: 1191

Sequence: 1 MYMEIFITGLLGASLLS.....INWVAVMTALIKMLMG 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	1191	100.0	236	9	US-09-738-626-6955
2	1191	100.0	236	14	US-10-156-142-8
3	1178	98.9	233	10	US-09-746-6608-52
4	1178	98.9	233	14	US-10-136-232-25
5	584	49.0	124	14	US-10-166-142-10
6	301	25.3	203	14	US-10-156-761-8659
7	133	11.2	224	14	US-10-156-761-7796
8	132.5	11.1	226	9	US-09-738-626-3665
9	128.5	10.8	212	12	US-10-374-9038-6
10	119	10.0	620	12	US-10-282-122A-77618
11	109.5	9.2	210	14	US-10-156-761-14994
12	103.5	8.7	181	12	US-10-282-122A-44873
13	98.5	8.3	217	14	US-10-156-761-10918
14	95	8.0	295	12	US-10-282-122A-77782
15	93.5	7.9	452	9	US-09-815-242-10183

16	93.5	7.9	452	12	US-10-282-122A-56572
17	93	7.8	211	12	US-10-282-122A-47826
18	93	7.8	459	9	US-09-738-626-6998
19	92.5	7.8	298	14	US-10-156-761-8730
20	91.5	7.7	446	9	US-09-351-780-2
21	91.5	7.7	446	9	US-09-738-626-6356
22	91.5	7.7	446	12	US-10-627-476-426
23	91.5	7.7	484	15	US-10-264-213-124
24	91.5	7.7	485	14	US-10-080-170-297
25	91	7.6	252	12	US-10-425-114-65374
26	91	7.6	361	14	US-10-156-761-9054
27	91	7.6	390	15	US-10-369-493-8216
28	91	7.6	624	9	US-09-738-626-4289
29	91	7.6	628	12	US-10-627-476-550
30	91	7.6	628	12	US-10-627-476-666
31	90.5	7.6	472	14	US-10-080-170-380
32	89.5	7.5	205	9	US-09-927-395-2
33	89.5	7.5	205	9	US-09-847-392-2
34	89.5	7.5	340	9	US-09-815-242-11980
35	89	7.5	477	15	US-10-369-493-18633
36	88.5	7.4	436	14	US-10-081-872-28
37	88.5	7.4	436	15	US-10-385-305-28
38	88	7.4	436	14	US-10-081-872-42
39	88	7.4	436	14	US-10-228-063-1
40	88	7.4	436	15	US-10-385-305-42
41	88	7.4	454	14	US-10-228-063-13
42	88	7.4	460	14	US-10-228-063-14
43	88	7.4	460	14	US-10-228-063-33
44	88	7.4	460	14	US-10-228-063-35
45	88	7.4	518	14	US-10-228-063-15

ALIGNMENTS

RESULT 1

US-09-738-626-6955
; Sequence 6955, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6955
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6955

Query Match 100.0%; Score 1191; DB 9; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.1e-117;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYMEIFITGLLGASLLSISGPQVIVIKQIKREGILAVLVCLISDVFLIAGTLGV 60

```

Db 1 MVMEIFITGLGASLLSIGPQNVIVKQIKREGLIAVLLVCLISDVFLIAGTLGV 60
QY 61 DLSNAAPIVLDIRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIEETPTVDDTPLGGS 120
Db 61 DLSNAAPIVLDIRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIEETPTVDDTPLGGS 120
QY 121 AVATDTRNRVRVSVVDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRW 180
Db 121 AVATDTRNRVRVSVVDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRW 180
QY 181 IFAGAFASLIWFFLVGFGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 236
Db 181 IFAGAFASLIWFFLVGFGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 236

RESULT 2
US-10-166-142-8
; Sequence 8, Application US/10166142
; Publication No. US20030124687A1
; GENERAL INFORMATION:
; APPLICANT: GUNJI, YOSHIYA
; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
; FILE REFERENCE: 223789US
; CURRENT APPLICATION NUMBER: US/10/166,142
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001-1777075
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-10-166-142-8

```

```

Query Match 100.0%; Score 1191; DB 14; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.1e-117;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVMEIFITGLGASLLSIGPQNVIVKQIKREGLIAVLLVCLISDVFLIAGTLGV 60
Db 1 MVMEIFITGLGASLLSIGPQNVIVKQIKREGLIAVLLVCLISDVFLIAGTLGV 60
QY 61 DLSNAAPIVLDIRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIEETPTVDDTPLGGS 120
Db 61 DLSNAAPIVLDIRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIEETPTVDDTPLGGS 120
QY 121 AVATDTRNRVRVSVVDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRW 180
Db 121 AVATDTRNRVRVSVVDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRW 180
QY 181 IFAGAFASLIWFFLVGFGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 236
Db 181 IFAGAFASLIWFFLVGFGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 236

RESULT 3
US-09-746-660A-52
; Sequence 52, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING

```

```

; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 52
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-746-660A-52

Query Match 98.9%; Score 1178; DB 10; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.4e-115;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEIFITGLGASLLSIGPQNVIVKQIKREGLIAVLLVCLISDVFLIAGTLGV 63
Db 1 MEIFITGLGASLLSIGPQNVIVKQIKREGLIAVLLVCLISDVFLIAGTLGV 60
QY 64 SNAAPIVLDIRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIEETPTVDDTPLGSSAVA 123
Db 61 SNAAPIVLDIRWGGIAYLLWFAVMAAKDAMTNKVEAPQIIEETPTVDDTPLGSSAVA 120
QY 124 TTRNRVRVSVVDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRWIFA 183
Db 121 TTRNRVRVSVVDKQVWVKPMLMAVLTLNPNAYLDAFVFIGVGAQYGDTRWIFA 180
QY 184 AGAFASLIWFFLVGFGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 236
Db 181 AGAFASLIWFFLVGFGAALSRPLSSPKVWRWVNVVAVMTALAIKMLMG 233

RESULT 4
US-10-196-232-25
; Sequence 25, Application US/10196232
; Publication No. US20030113899A1
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, MIKIKO
; APPLICANT: ITO, HISAO
; APPLICANT: GUNJI, YOSHIYA
; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: 225391USO
; CURRENT APPLICATION NUMBER: US/10/196,232
; CURRENT FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: JP 2001-224586
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-196-232-25

Query Match 98.9%; Score 1178; DB 14; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.4e-115;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 MEIFITGLLGASLLISGPNQVNLVIKQIKREGLIAVLLVCLISDVFLFIAGTLGVDLL 63
Db 1 MEIFITGLLGASLLISGPNQVNLVIKQIKREGLIAVLLVCLISDVFLFIAGTLGVDLL 60
QY 64 SNAPIVLDIMRWGGIAYLLWFVAAKADMTNKVEAPQIIETEPTVPDDTPLGGSNAVA 123
Db 61 SNAPIVLDIMRWGGIAYLLWFVAAKADMTNKVEAPQIIETEPTVPDDTPLGGSNAVA 120
QY 124 TDRNRVRVSVSKORVWVKPMLMAIVLTWLPNAYLDARFVIGGVGAQYGDGTGRWIFA 183
Db 121 TDRNRVRVSVSKORVWVKPMLMAIVLTWLPNAYLDARFVIGGVGAQYGDGTGRWIFA 180
QY 184 AGAFAASLIWFLVPGGAALSRPLSSPKWRWVNVVAVMTALAIKMLMG 236
Db 181 AGAFAASLIWFLVPGGAALSRPLSSPKWRWVNVVAVMTALAIKMLMG 233

RESULT 5
US-10-166-142-10
; Sequence 10, Application US/10166142
; Publication No. US20030124687A1
; GENERAL INFORMATION:
; APPLICANT: GUNJI, YOSHIYA
; APPLICANT: YASUEDA, HISASHI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
; FILE REFERENCE: 223789US
; CURRENT APPLICATION NUMBER: US/10/166,142
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001-1777075
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-10-166-142-10

Query Match 49.0%; Score 584; DB 14; Length 124;
Best Local Similarity 100.0%; Pred. No. 2e-53;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYMEIFITGLLGASLLISGPNQVNLVIKQIKREGLIAVLLVCLISDVFLFIAGTLGV 60
Db 1 MYMEIFITGLLGASLLISGPNQVNLVIKQIKREGLIAVLLVCLISDVFLFIAGTLGV 60
QY 61 DLLSNAAPIVLDIMRWGGIAYLLWFVAAKADMTNKVEAPQIIETEPTVPDDTPLG 118
Db 61 DLLSNAAPIVLDIMRWGGIAYLLWFVAAKADMTNKVEAPQIIETEPTVPDDTPLG 118

```

```

; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8659
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8659

Query Match 25.3%; Score 301; DB 14; Length 203;
Best Local Similarity 32.1%; Pred. No. 2.3e-23;
Matches 72; Conservative 32; Mismatches 86; Indels 34; Gaps 3;

QY 10 GLLLGASLLISGPNQVNLVIKQIKREGLIAVLLVCLISDVFLFIAGTLGVDLLSNAPI 69
Db 11 GFETGLSLIVAIQNAFVLRQGIERRDAVLAIVGICALLIALLGALGVGVAVVWVAMR 70
QY 70 VLDIMRWGGIAYLLWFVAAKADMTNKVEAPQIIETEPTVPDDTPLGGSNAVATDTRNR 129
Db 71 ALTAVALVGGAFLLVYGALARRVF-----RPAURAGCAPESR 110
QY 130 VRVEVSVDKORVWVKPMLMAIVLTWLPNAYLDARFVIGGVGAQYGDGTGRWIFAAGAF 189
Db 111 RRA-----VLTCLALTWLPNPHVVDVTVFLGSLIADRGSL-RWTEGLGAALA 156
QY 190 SLIMEPLVPGGAALSRPLSSPKWRWVNVVAVMTALAIKLM 233
Db 157 SLCWFAALGFGSLRGLRFLARPSAWRVDLDAVVAATMLTWGATLI 200

RESULT 7
US-10-156-761-7796
; Sequence 7796, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7796
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7796

Query Match 11.2%; Score 133; DB 14; Length 224;
Best Local Similarity 24.2%; Pred. No. 1.3e-05;
Matches 62; Conservative 40; Mismatches 92; Indels 62; Gaps 15;

QY 4 MEIFITGLLGASLLISGPNQVNLVIKQIKREGLIAVLLVCLISDVFLFIAGTLGV 55
Db 1 MSVDIVG-FLGVVIVVYVPGDFLVVRSATEHPAKRAALGAQSGUCV-----HMLA 54
QY 56 GTLVGDLISNAAPIVLDIMRWGGIAYLLWF-----AVMAKADMTNKVEAPQII-----EETE 108
Db 55 AAVGLSLIAARSNAVYDAIRLUGAAYLVILGVRAVLAARRAARRAAGREAVGVGDEGTD 114
QY 109 PTVP-DUTPLGGSNAVATDTRNRVRVSVSKORVWVKPMLMAIVLTWLPNAYLDARF 167
Db 115 PRTPPEEAPRGR-----WRSQGTQGTFLTNVLPNKAAL-----FF 149
QY 168 GGVGAQY-----GDTGRWIFAAGAFASL-----IWFPLVGVGAALSRPLSSPKV-----WR-- 215

```

Db 150 LSLPQVHGGSSTQIPFLGTLDIVGVAYFALVAV-AARLRAFARPKVRHGWEIT 208
 QY 216 --WINVVAVMTALA 229
 Db 209 TGLWFTAIGGVAAA 224

RESULT 8
 US-09-738-626-3665
 ; Sequence 3665, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIALI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 3665
 ; LENGTH: 226
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; US-09-738-626-3665

Query Match 11.1%; Score 132.5; DB 9; Length 226;
 Best Local Similarity 24.2%; Pred. No. 1.5e-05;
 Matches 58; Conservative 45; Mismatches 98; Indels 39; Gaps 12;
 QY 7 FRTGLGASLLSTGPNVLIKQIK--REGLIAVLLVCLISDVFLTAGTLGVDLLS 64
 Db 6 FLALFLVAAATSPGDLQIILSAKRRDGLVTAVGLWGNISWI--IASLLGLSALI 64
 QY 65 NAAPVLDIMRWGGIAYLLFVAVMAAKDMTKVKEAPQIIETEPTVDDTPLGGSVAT 124
 Db 65 STYPALINLLQVGGVLTWVGICAVRSWTKR-----STQQAADSQAVENTLVTA 116
 QY 125 DTRNRVREVSVDKQVRVWVPMALVLTWL--NENAYLDAFVFIGVGAQY--GDTG-RW 180
 Db 117 TMAIS-----VGVPAIRSGIATNLSNPAVL-----FFGSVFAQVREPDGIGW 160
 QY 181 IFAGAF--AASLIWFFLVGGAALSRLPSS--PKVWRWVAVVMTALAIKMLMG 236
 Db 161 SIFIGVFLTLGLWF--VGF--AVLVKLAAGLTNGAIDLLTGVIFIGLGMFIPEG 216

RESULT 9
 US-10-374-903A-6
 ; Sequence 6, Application US/10374903A
 ; Publication No. US20040038250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Oviedo
 ; APPLICANT: Astur Pharma, S.A.
 ; TITLE OF INVENTION: The gene cluster for thienamycin biosynthesis,
 ; FILE REFERENCE: Thienamycin-UG-AP
 ; CURRENT APPLICATION NUMBER: US/10/374,903A
 ; CURRENT FILING DATE: 2003-02-26

; NUMBER OF SEQ ID NOS: 33
 ; SEQ ID NO 6
 ; LENGTH: 212
 ; TYPE: PRT
 ; ORGANISM: Streptomyces cattleya
 ; US-10-374-903A-6
 Query Match 10.8%; Score 128.5; DB 12; Length 212;
 Best Local Similarity 25.0%; Pred. No. 3.5e-05;
 Matches 62; Conservative 34; Mismatches 93; Indels 59; Gaps 12;
 QY 4 MEITITGLL--LGASLLLSI--QPQNVLVKQGI--KREGLIAVLLVCLISDVFLF-IAG 56
 Db 1 MEPMLTALAFGLGACVLIAAAGPSTMLIIHQSLHSRRAGFLTVL--GNETGVLTWGVA 58
 QY 57 TLGVOLLSNAAPIVLIDIMRWGGIAYLLFVAVMAAKDMTKVKEAPQIIETEPTVDDTP 116
 Db 59 ALGTLTALLAASRTAYDVRIGGAVLVWVYGVQTURAARG-----EAFSAADD-- 107
 QY 117 LGGSVATDTRNRVREVSVDKQVRVWVPMALVLTWLNPAYLDAFVFIGVGAQYGD 176
 Db 108 -----EAAVPRSGW-KIYRSGLLLNLANPKAAVFAVMSFL----PQFVP 146
 QY 177 TGR-----WIFAGAPAA-----SLIWFPLVGFGAALSRLSPKVRWVWVAVV 224
 Db 147 AGAPKLPVITAAAPQALFEVGYGYGVYVWF-----VGRMKRVISRAGVRRRLREQVSGGV 200
 QY 225 MTALAIAKL 232
 Db 201 LVLLGIRM 208
 RESULT 10
 US-10-282-122A-77618
 ; Sequence 77618, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 77618
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77618

Query Match      10.0%; Score 119; DB 12; Length 620;
Best Local Similarity 22.8%; Pred. No. 0.0015;
Matches 63; Conservative 29; Mismatches 72; Indels 112; Gaps 14;

QY 2 VIMEIFITGLLGSALLSISGPNVLVIKQIKKEGLIA-----VLLVCLISDVFLFI 54
Db 102 MLLGILATGFFGGAIVTP--AMSVLSAVEGI-----AAVQPLAPYVLPIDAMIIIVAUFA 155
QY 55 AGTLGVGLISN-AAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETEPTVD 113
Db 156 VQAMGTERTIGRFFAPVM-----LLWFLVLAL----- 181
QY 114 DTPGGSAVADTRNRVRVEVSVDKORVWVKPMMLMAIVLTWLNPNAYLDAPVF----- 166
Db 182 ---LGAHA-----IWHAPQ---VLRALNP-AVAVHFVLLYQHTL 214
QY 167 -----IGVGAQYDGTGRWIFAGAFPAASLIWFLPV-----GFGA--AALSR 206
Db 215 FIUGLVLSVGTVALYADMGHF---GIKPIRIAPFALWPSLLNLYFGQAYLLTISA 270
QY 207 PLSS-----PKVRWNVVAVVMTALAIKMLMG 236
Db 271 PTGSTFSLAFKAMLPILLATFATVIASQAVISG 306

RESULT 11
US-10-156-761-14994
; Sequence 14994, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14994
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14994

Query Match      9.2%; Score 109.5; DB 14; Length 210;
Best Local Similarity 22.7%; Pred. No. 0.0035;
Matches 53; Conservative 36; Mismatches 107; Indels 37; Gaps 7;

QY 8 ITGLLGALLSISGPNVLVIKQIK---REGLIAVLLVCLISDVFLFIAGTLGVGLLS 64
Db 6 VAGLAGYGIAPVGAATVYLSLTARTSLRTGYCAALGVATADGLVALVA-ALGGSALA 64
QY 65 NAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAPQIIEETEPTVDTPGGSAVAT 124
Db 65 AALQPVLPWRWASGLVIALAVRGATAAVERH-YRSPRAHAAARETPSP----- 113
```

```
QY 125 DTRNRVRVEVSVDKORVWVKPMMLMAIVLTWLNPNAA--YLDAFVFTGGVGAQ-YGDTGRWI 181
Db 114 -----ARAYLGLLGIITLLNPTVTYFAALV-LGSRAQAQVLPLEQGV 154
QY 182 FAAGAFPAASLIWFLPVGFGAAALSRPLSSPKVWRWNVVAVVMTALAIKML 234
Db 155 FVLAAFVASASQVLLAGGALLGRALTGRGRRLVTALLSSAVILGLAVRLM 207

RESULT 12
US-10-282-122A-44873
; Sequence 44873, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 44873
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-44873

Query Match      8.7%; Score 103.5; DB 12; Length 181;
Best Local Similarity 24.1%; Pred. No. 0.012;
Matches 48; Conservative 35; Mismatches 65; Indels 51; Gaps 10;

QY 42 LLVCLISDVFLFIAGTLGVGLLSNAAPIVLDIMRWGGIAYLLWFVMAAKDAMTNKVEAP 101
Db 19 LQALIQIMIVAA---GVGLFATTPLAFOAVKMFVAYLLYLA----- 60
QY 102 QIIEETEPTVDTPGGSAVATDTRNRVRVEVSVDKORVWVKPMML-AIVLTWLNPNAY 160
Db 61 -YQWTAP-VKD-----LEIQHEKKOKSVSALLNGFVYVNIENPK-- 98
QY 161 LDAFVFTGGVGAQYDGTGR--WIFAGAFPAASLIWFLPV---GFG--AALSRPLSPKV 213
```

Db 99 --ATVELLAVLPOFLDLKSPQIYOL-IMAATVMTVIDLIWAGYTGKASKVLRLRLSPKQ 155
 QY 214 WRWNVVAVVMTALAUKL 232
 Db 156 QKYLNGRFVWFSCAALLL 174

RESULT 13

US-10-156-761-10918
 ; Sequence 10918, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; PRIOR FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 10918
 ; LENGTH: 217
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-10918

Query Match 8.3%; Score 98.5; DB 14; Length 217;

Best Local Similarity 21.6%; Pred. No. 0.052; 93; Indels 73; Gaps 12;
 Matches 55; Conservative 34; Mismatches 93;

QY 7 FITGLLGASLLISGPQNLVI-----KQIKREGGLAVLLVCLISDVFLTAGTLGVLD 62
 Db 10 YIAGLVL---IVLLFGPNSLVLSVAARKGVRAGYTAAAGWVC--GDTVMTLSAAGVAS 64
 QY 63 LSNAAPIVLDIMRWGGIAYLWFAVMAKDA-----MTNKEAPQIIEETETPTVD 113
 Db 65 LIQANALIFGIVKAGYGLTFLAFGLRAAEMWRTRRADAAPV----- 113
 QY 114 DPLGSSAVATDTRNRVREVSVDKQVWVKPMLMAIVLTWLPNAYLDAFVFGVGAQ 173
 Db 114 -APAGE-----RFRRAFVSLFNPKALFFVAFV-----VQ 145
 QY 174 YDGTGRWIFAA-----GAFV--ASLIWFLVFGF-----AAALSRLSPKPKVWRINVV 221
 Db 146 FVDFG-YAYFALSFFVLGAFQAQLASFLYLTALIFSGTKLAAAFR---PKRLSAGATTAA 201
 QY 222 AVVMTALAIKLMJM 236
 Db 202 GALEFLGFAVKLTLAG 216

RESULT 14

US-10-282-122A-77782
 ; Sequence 77782, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 77782
 ; LENGTH: 295
 ; TYPE: PRT
 ; ORGANISM: Versinia pestis
 US-10-282-122A-77782

Query Match 8.0%; Score 95; DB 12; Length 295;

Best Local Similarity 24.2%; Pred. No. 0.118;
 Matches 63; Conservative 42; Mismatches 85; Indels 70; Gaps 13;

QY 1 WYMEIFITGLLGASLLISGPQNLVIKQIK--REGGLAVLLVCLISDVFL-FIGT 57
 Db 16 ILLIVIMISIQSGASLAKSLFP---LVGAGQITSRLGIGTLLIFVIFKPKWRFPAAG- 71
 QY 58 LGVDLLSNAAPIVLDIMRWGGIAYLWFAVMAKDAKMTNKVEAPQIIEETETPTVD 117
 Db 72 -----SRPLLIYGVAGLGMFLFVMSL-----KTVP----- 98
 QY 118 GGSVATDTRNRVREVSVDKQV---WVKPMLMAIVLTW-LNPNAY---LDAFVFIGG 169
 Db 99 LGIAVALEFTGPLAVAMFSSRAVDFTWVG---LAILGLMFLPLGHTGTIDLFGAACA 155
 QY 170 VCA-----QVDTGRWIFAAAGAFASLIWFLP-VGFGAAALSRLSPKPKVWR 215
 Db 156 LGAGACWAIYISGQXAGDHPGTVAVGSLTALIFCPIGVANGMALFDPAILP----- 211
 QY 216 WINVVAVVMTALAIKLMJM 235
 Db 212 -IALAVATLSTALPYSLEWV 230

RESULT 15

US-09-815-242-10183
 ; Sequence 10183, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE OF INVENTION: Prokaryotes
 FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10183

LENGTH: 452

TYPE: PRT

ORGANISM: Escherichia coli

US-09-815-242-10183

Query Match 7.9%; Score 93.5; DB 9; Length 452;

Best Local Similarity 19.4%; Pred. No. 0.47;

Matches 52; Conservative 44; Mismatches 95; Indels 77; Gaps 14;

QY 8 ITGLLG--ASLLSIGPQNVLV-----IKQGIKR--EGLIAVLVCLISDVFLFIAGT 57
 Db 150 ISGIIGLVVAEMFTFVRNWKLPDVPASVSFSALIPGFIILSVNGIIAWALNT 209
 QY 58 LG-----VDLLSNAPIVLDIRWGGIAY--LLWFAM---AAKAMTNKVEAPQIIE 105
 Db 210 WGTNPHQIIMDTISTPLASIGSVGVAWVIFVFLWFFGIHGALALTALDNGIMTPWALE 269
 QY 106 ETEPTVDDTFLGGSNAVAT-DTRNRVRVSVVDKQ-RVVVKPMLMAVLVTLNPNAYLDA 163
 Db 270 N-----IATYQYGSVEAALAGAKTFHIWAKPM-----LDS 300
 QY 164 FVFIGGVG-----AOYGDGRWIFAGAFASLLWPELVGFGAALSRL 208
 Db 301 FIFLGGGAILGLILAIFFASRRADYQVAKLALPSGIFQ---INEFIL-FGLPIIMNPV 356
 QY 209 SSPKVWRWVNVVAVMTALAIKLMNG 236
 Db 357 MEIPE-----VLVQPIILAAITLAAAYNG 379

Search completed: March 18, 2004, 06:44:48
 Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:20:36 ; Search time 20 Seconds
(without alignments)
1135.060 Million cell updates/sec

Title: US-09-105-117K-2
Perfect score: 1191
Sequence: 1 MYMEIFITGLLGASLLLS.....INNVAVVMTALAIKLMIMG 236

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl.*
2: Pirl.*
3: Pirl.*
4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	350	29.4	211	1 Q0EC5A	hypothetical 23K p
2	347	29.1	211	2 AH0874	probable membrane
3	343	28.8	211	2 B91103	hypothetical prote
4	340	28.5	211	2 B85948	hypothetical prote
5	335	28.1	205	2 AG0112	probable LysE type
6	332	27.9	200	2 D83100	probable transport
7	326.5	27.4	202	2 AG2690	LysE family transp
8	326.5	27.4	202	2 C97472	probable membrane
9	319	26.8	201	2 C70744	hypothetical prote
10	318	26.7	199	2 H70756	hypothetical prote
11	302.5	25.4	204	2 AD3411	transporter, lysE
12	245.5	20.6	211	2 B83218	LysE/YggA family p
13	238	20.0	211	2 S57940	YggA protein homol
14	176.5	14.8	220	2 D69838	conserved hypotet
15	173.5	14.6	210	2 F64609	conserved hypotet
16	170.5	14.3	210	2 B71906	hypothetical prote
17	164	13.8	205	2 B89854	conserved hypotet
18	154.5	13.0	205	2 C97219	uncharacterized co
19	153.5	12.9	213	2 E97789	hypothetical prote
20	151.5	12.7	208	1 B69066	conserved hypotet
21	148	12.4	200	2 G83703	hypothetical prote
22	136.5	11.5	213	2 F83444	hypothetical prote
23	132	11.1	210	2 E87252	efflux protein, Ly
24	130	10.9	210	2 D84016	hypothetical prote
25	120.5	10.1	208	2 G87305	efflux protein, Ly
26	119	10.0	620	2 F82449	potassium uptake p
27	117.5	9.9	197	2 B83280	hypothetical prote
28	115.5	9.7	216	2 AH3203	RhtB family transp
29	114.5	9.6	208	2 G84086	dihydrodipicolinat

30	113	9.5	222	2 E87264	efflux protein, Ly
31	113	9.5	235	2 AD2904	RhtB family transp
32	113	9.5	235	2 F97679	hypothetical prote
33	111	9.3	212	2 AD0714	probable membrane
34	109.5	9.2	218	2 A96009	probable amino aci
35	109.5	9.2	450	2 E72615	probable hlyE AP81
36	108.5	9.1	203	2 AB2683	RhtB family transp
37	108.5	9.1	239	2 H97464	conserved hypotet
38	107.5	9.0	284	2 H69277	branched-chain ami
39	107	9.0	210	2 G83082	hypothetical prote
40	105	8.8	206	2 A98223	theonine efflux p
41	104.5	8.8	349	2 AG3010	hypothetical prote
42	104.5	8.8	349	2 G98273	iron(III) dictrat
43	104	8.7	205	2 B83356	conserved hypotet
44	102	8.6	206	2 AH0466	theonine efflux p
45	102	8.6	212	2 T43921	yfud protein [impo

ALIGNMENTS

RESULT 1

Q0EC5A
hypothetical 23K protein (sbm-fba intergenic region) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 31-Dec-1990 #sequence_revision 17-Oct-1997 #text change 01-Mar-2002
C:Accession: B65077; S04736
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B65077
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-211 <BLAT>
A:Cross-references: GB:AE000375; GB:U00096; NID:gl789282; PIDN:AACT5960.1; PID:gl789290.
A:Experimental source: strain K-12, substrain MG1655
R:Alfred, P.R.; Perham, R.N.
Mol. Microbiol. 3, 723-732, 1989
A:Title: Identification, molecular cloning and sequence analysis of a gene cluster encod
hyde 3-phosphate dehydrogenase of Escherichia coli.
A:Reference number: S04730; MUID:89313302; PMID:2546007
A:Accession: S04736
A:Molecule type: DNA
A:Residues: 15-211 <ALF>
A:Cross-references: EMBL:X14436; NID:g41417; PIDN:CAA32607.1; PID:g41425
C:Genetics:
A:Gene: yggA
A:Map position: 63 min
C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)

Query Match	29.4%	Score 350;	DB 1;	Length 211;
Best Local Similarity	35.7%	Pred. No. 1.5e-24;		
Matches	81;	Conservative	37;	Mismatches 75; Indels 34; Gaps 4;
Qy	7	FTGTLGASLLLSIGQNVVVKQIKREGILAVLVCLISDVFLFIAGTGLVDLSNA	66	
Db	5	YFQGLAAGAMTILPGQNAFVNQGIQIRQYHIMALLCAISDLVLCAGIFGGSALLMQ	64	
Qy	67	APVLDIMWGGIATLLWFAVMAKDMTKVAPQIIEETPTVPDDTPLGGSATVDT	126	
Db	65	SPMLALVTWGGVAFPLWYGFAGFTAMSSNIEI	98	
Qy	127	RNRVRVEVDKQRVWVKPMALVILTWLNPNAYIDAFVFIGVGQAQYG-DTGRWIFAAG	185	
Db	99	-----ASAEMWQGRW-KIATMTLAVTWLNPHVYLDFTFVLGSLGQLDVEPKRW-FALG	151	
Qy	186	AFASLIWFLVFGGAALSRPLSSPKVRWVNVVAVVMTALAIKL	232	
Db	152	TISAGFLNFFGLLAAWLAAPRLTAKAQRIINLVGCVNWFIALQL	198	

Cy **Dd**

67 APVLIDIMRWGGIAYLLIFANWAAKADMTNKVEAPQIIETEPTVPDDTPLGSSAVATDT 126
 :
 :
 :
 :
 :
98

65 SPMLIALVTWGVSFLWTGFGAFKTAMSNIEL-----
 :
 :
 :
 :
 :

Cy **Dd**

127 RNRVRVESVDQRVWKVKMPLMAIVLTWLNPYALDAFVFVGGAQG-DTGWRIFAAG 185
 :
 :
 :
 :
 :
99 -----ASAENVLKGRW-KTIATMLAVTWLPHYLTDFTLVGLSGQLDVBPKRW-FALG 151
 :
 :
 :
 :
 :

Cy **Dd**

186 AFASLIWFPLGFCAALSRLSPSKVPRVINNVVAVMATAIKL 232
 :
 :
 :
 :
 :
152 TISASFLEFFGLAILAAMPLRTAKSQRIINLVGCVMWFIAQL 198
 :
 :
 :
 :
 :

RESULT 4
E85948
hypothetical protein yggA [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: E85948
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grobeck, E.J.; Davis, A.M.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: AB54480; MUID:21074933; PMID:11206551
A;Accession: E85948
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <STO>
A:Cross-references: GB:A5005174; NID:g12517455; PIDN:AAGS8049.1; GSPPB:GN00145; UMGF:P4
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A:Gene: YGG
C;Superfamily: Escherichia coli hypothetical 23k protein (sbm-fba intergenic region)

Query Match 28.5%; Score 340; DB 2; Length 211;
Best Local Similarity 34.5%; Pred. No. 1-2e+23;
Matches 79; Conservative 37; Mismatches 79; Indels 32; Gaps 2;

Cy **Dd**

7 FITGLLGASILSIGPNNVIKGIRKLGIALLVCLSDVLFPIAGTLGVDLISA 66
 :
 :
 :
 :
 :
5 YFOGLAAGAAMILPGPFNAFMNQRRQHIMIALLCAISDVLCIGAFGGSAALLMQ 64
 :
 :
 :
 :
 :

Cy **Dd**

67 APVLIDIMRWGGIAYLLIFANWAAKADMTNKVEAPQIIETEPTVPDDTPLGSSAVATDT 126
 :
 :
 :
 :
 :
65 SPMLIALVTWGVSFLWTGFGAFKTAMSNIEL-----
 :
 :
 :
 :
 :

Cy **Dd**

127 RNRVRVESVDQRVWKVKMPLMAIVLTWLNPYALDAFVFVGGAQG-DTGWRIFAAG 186
 :
 :
 :
 :
 :
99 -----ASAENVLKGRW-KTIATMLAVTWLPHYLTDFTLVGLSGQLDVBPKRWFAALTGT 152
 :
 :
 :
 :
 :

Cy **Dd**

187 AFASLIWFPLGFCAALSRLSPSKVPRVINNVVAVMATAIKL 232
 :
 :
 :
 :
 :
153 ISASFLEFFGLAILAAMPLRTAKSQRIINLVGCVMWFIAQL 198
 :
 :
 :
 :
 :

RESULT 5
AG0112
probable lysE type translocator YPO0918 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C;Accession: AG0112
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.E
deno-Parrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.; I
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AG0112
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-205 <CUR>
A:Cross-references: GB:AL590842; PIDN:CAC99762.1; PID:g15978989; GSPPB:GN00175
C;Genetics:

QY 64 SNAAPVILDMRGGIAYLLWFAVMAKADMTNKVEAPQIIEETPTVDDTPLGGSVA 123
 Db 63 SAIMPALDPIRMVAGAAFLWYGAKSLSYALRSS-EVLSVAERREAS----- 108
 QY 124 TDTNRNRVSVSKQKRVVWVPMALVILWLNPNAYLDADFVIGVGAGYQDGTGRWIFA 183
 Db 109 -----LW-QTALCALTFPLNPHVILDTVVLGTISTQTFQFK-TFA 149
 QY 184 AGAFAASLWPLVGFCAALSRPLSSPKVWRWVNVVAVMTALAIAKL 234
 Db 150 AGAATGSLLEFFSGYGARWLPIPEKESAWRILEGVIAITWMAIAFLWV 200
 RESULT 9
 C70744
 hypothetical protein Rv0488 - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: C70744
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: C70744
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-201 <COL>
 A:Cross-references: GB:Z77162; GB:AL123456; NID:g3261606; PIDN:CAB00949.1; PID:g1449286
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv0488
 C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)
 Query Match 26.8%; Score 319; DB 2; Length 201;
 Best Local Similarity 32.4%; Pred. No. 9.6e-22;
 Matches 71; Conservative 50; Mismatches 64; Indels 34; Gaps 4;
 QY 16 SLLSISGPQNVLIKQIKREGIIVLVCLISDVFLFTAGTIGVLLSNAAPVILDMR 75
 Db 3 TLAVAGPQNAFLVRQGRREYVIVLALGADGALIAAGVGGFAALIAHNNMTLVAR 62
 QY 76 WGIAYLLWFAVMAKADMTNKVEAPQIIEETPTVDDTPLGGSVAATNRVREVS 135
 Db 63 FGAAPFLIGYALLAARNW-----RPSGLVPESS--GPAAL----- 96
 QY 136 VDKQKRVVWVPMALVILWLNPNAYLDADFVIGVGAGYQDGTGRWIFAASLIWFP 195
 Db 97 -----IGVQMCILVTFNPHVILDTVVLIGALANESDL-RWFFGAGNAAVWFA 148
 QY 196 LVGFCAALSRPLSSPKVWRWVNVVAVMTALAIAKL 234
 Db 149 VLGSAGRIQPFATPAANRILDALVAVTMIGVAVVVLV 187
 RESULT 10
 H70756
 hypothetical protein Rv1986 - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: H70756
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: H70756
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A:Residues: 1-199 <COL>
 A:Cross-references: GB:Z74025; GB:AL123456; NID:g3261586; PIDN:CAA98398.1; PID:g1403469
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv1986
 C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)
 Query Match 26.7%; Score 318; DB 2; Length 199;
 Best Local Similarity 31.1%; Pred. No. 1.2e-21;
 Matches 70; Conservative 44; Mismatches 77; Indels 34; Gaps 3;
 QY 8 ITGALLGASLLISGPQNVLIKQIKREGIIVLVCLISDVFLFTAGTIGVLLSNA 67
 Db 6 VVGFLACFTLIALGAQNAFVLRQIGRHFVLPVVALCTVSDVILANGTAGFALIGH 65
 QY 68 PIVLDMRGGIAYLLWFAVMAKADMTNKVEAPQIIEETPTVDDTPLGGSVAATDR 127
 Db 66 PRALNVVFGGAFLIGYGLAARAWRPVALIP-----SGATPVR 106
 QY 128 NRVREVSVDKQKRVVWVPMALVILWLNPNAYLDADFVIGVGAGYQDGTGRWIFAAGAF 187
 Db 107 -----LAEVIVTCAFTFNLPHVILDTVVLIGALANESDQ-RMLFGIGAV 151
 QY 188 AASLIWFPFLVGFCAALSRPLSSPKVWRWVNVVAVMTALAIAKL 232
 Db 152 TASAVWFATLGFAGRLRGLFTNPGSWRLDGLIIVVWVALGISL 196
 RESULT 11
 AD3411
 transporter, lysE family BME11274 [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
 C:Accession: AD3411
 R;DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujier, C.; Los, T.; Ivanova, M.; Mazur, M.; Gotsman, E.; Seikov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letes, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AD3411
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-204 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AAL52455.1; PID:g17983261; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME11274
 A:Map position: 1
 C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)
 Query Match 25.4%; Score 302.5; DB 2; Length 204;
 Best Local Similarity 30.2%; Pred. No. 3.1e-20;
 Matches 68; Conservative 43; Mismatches 81; Indels 33; Gaps 3;
 QY 8 ITGALLGASLLISGPQNVLIKQIKREGIIVLVCLISDVFLFTAGTIGVLLSNA 67
 Db 6 LSGFLGASLIITAGNAQNAFVLRQIGRHFVILCLICALSALLISAGVAGLTIAQS 65
 QY 68 PIVLDMRGGIAYLLWFAVMAKADMTNKVEAPQIIEETPTVDDTPLGGSVAATDR 127
 Db 66 PKLIAFTVLGAAFLWYASVAFRAFP--HPEAMQVKSNGAVSLK-----AAAAAT--- 113
 QY 128 NRVREVSVDKQKRVVWVPMALVILWLNPNAYLDADFVIGVGAGYQDGTGRWIFAAGAF 187
 Db 114 -----CLALTFLNPHVILDTVVLIGSLSARLSGPARAYGAA 152
 QY 188 AASLIWFPFLVGFCAALSRPLSSPKVWRWVNVVAVMTALAIAKL 232
 Db 153 TASCIFWFFALGYARLLQPIFAKPAARVLDLCIIGVWVALGISL 197
 RESULT 12
 B82318

Lyse/YggA family protein VC0481 [imported] - Vibrio cholerae (strain Ni6961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82318
R:Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Yamathavan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82318
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <H2I>
A:Cross-references: GB:AE004134; GB:AE003852; NID:95654900; PIDN:AAF93654.1; GSDB:GN001
A:Experimental source: serogroup O1; strain Ni6961; biotype El Tor
C:Genetics:
A:Gene: VC0481
A:Map position: 1
C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)

Query Match 20.6%; Score 245.5; DB 2; Length 211;
Best Local Similarity 28.6%; Pred. No. 5e-15; Mismatches 34; Indels 31; Gaps 3;
Matches 65; Conservative 34

Qy 6 IFITGLGASLLSIGPQNVLIKQIKREGIAVLVCLISDVFLTAGTLGVLLSN 65
Db 5 ILLQGSIGATMIPIGAGVAVLNQIKKHHLTTATCGVLDIMFIFLIGFGGALIS 64

Qy 66 AAPVILDMRWGGIAYLLFAVMAAKDAMTKVAPQIIEETPTVDDTPLGGSVATD 125
Db 65 QNTSLILGVTLLGFLCIGYFLSRAAL-----KPPQASESTA-----NPMAGRKAV- 113

Qy 126 TNRVRVSVSDQRVWVKPMALVLTWNPNAVLDAPFVIGVGAQYGDGTGWIIPAAG 185
Db 114 -----IFGAPVTVFNPHLYLDTVTLGSGGQGGDERISFAIG 153

Qy 186 AFAASLIWFLVFGGAALSRPLSSPKVWKNVWVAVVWTPALAIKL 232
Db 154 TILASFVWFFTLISLGAALKSTLSKPRVRQVIDMAVAMVFIAPAL 200

RESULT 13
S57940
YggA protein homolog - Aeromonas hydrophila
C:Species: Aeromonas hydrophila
C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
C:Accession: S57940
R:Swift, S.; Fish, L.; Williams, P.; Stewart, G.S.A.B.
submitted to the EMBL Data Library, July 1995
A:Reference number: S57938
A:Accession: S57940
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <SWI>
A:Cross-references: EMBL:X89469; NID:9899143; PIDN:CAA61655.1; PID:9899146

Query Match 20.0%; Score 238; DB 2; Length 211;
Best Local Similarity 26.6%; Pred. No. 2.4e-14; Mismatches 36; Indels 54; Gaps 4;
Matches 58; Conservative 36

Qy 17 LLSLSIGPQNVLIKQIKREGIAVLVCLISDVFLTAGTLGVLLSNAAPVILDMRW 76
Db 1 MIIPGAGNAFVLSRGHNRHLLTATLCLDLVIGVFGGANLLAASPGLALLTW 60

Qy 77 GGIAYLLFAVMAAKDAMTK-----VAPQIIEETPTVDDTPLGGSVATDTRVRV 132
Db 61 GGVLFLGFGHGRSLRSGAGQAKLADSPQL-----MGYKSV----- 97

Qy 133 EVSVDKQRVWVKPMALVLTWNPNAVLDAPFVIGVGAQYGDGTGWIIPAAGASLI 192
Db 98 -----LMTIGVTLNPHLYLDTLMLGSGSQPAEELRSFAAVAMLASLV 144

Qy 193 WFLPVFGAALSRPLSSPK-----VWRW 216
Db 145 WFLVAFGAVLVPLARSRQGVSKLLILLVSPCWGW 182

RESULT 14
D69838
conserved hypothetical protein yisU - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: D69838
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beretti
C.; Bron, S.; Brouillet, S.; Bruchci, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chk
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Konigstein, G.; Krogth, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
keuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, I
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D69838
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-220 <KUN>
A:Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12926.1; PID:G2633422
A:Experimental source: strain 168
C:Genetics:
A:Gene: yisU
C:Superfamily: Escherichia coli hypothetical 23K protein (sbm-fba intergenic region)

Query Match 14.8%; Score 176.5; DB 2; Length 220;
Best Local Similarity 22.9%; Pred. No. 1e-08; Mismatches 53; Conservative 36; Indels 43; Gaps 4;
Matches 53

Qy 4 MEIFITGLGASLLSIGPQNVLIKQIKREGI---IYVLLVCLISDVFLTAGTLGV 60
Db 18 MNAIHGIVLAFGLILPLGVQNVFFQAGLQKHWRALPAVISAASVCDTLILVAVAGV 77

Qy 61 DLLSNAAPVILDMRWGGIAYLLFAVMAAKDAMTKVAPQIIEETPTVDDTPLGGS 120
Db 78 SVIVQELPFVETVMAGGFLILNGW----- 105

Qy 121 AVATDTRNRVREVSVDQRVWV---KPLMALVLTWNPNAVLDAPFVIGVGAQYGDGTG 178
Db 106 -----TWN-IRPNTSQNKHTFTPKQAFAAAVSLNPHALDITGIVGTSLQVSGLE 159

Qy 179 RUIFAAGAAASLIWFLVFGGAALSRPLSSPKVWKNVWVAVVWTPALAIKL 229
Db 160 KMLFMAACIAVSWITFISLAIAGRLFTIDTISGRMLIVNKCSAAVWMAAA 210

RESULT 15
F64609
conserved hypothetical integral membrane protein HP0718 - Helicobacter pylori (strain 26
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 17-Mar-2000
C:Accession: F64609
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: F64609
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

Search completed: March 12, 2004, 18:26:47
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:04:50 ; Search time 18 Seconds
(without alignments)
682.698 Million cell updates/sec

Title: US-09-105-117K-2

Perfect score: 1191

Sequence: 1 MWMEIFINGLIGASLLS.....INVVAVVMTALAIKMLMG 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1178	98.9	233	1 LYSE CORGL	P94533 corynebacte
2	864.5	72.6	228	1 LYSE CORGL	Q8qtm4 corynebacte
3	350	29.4	211	1 YGGA_ECOLI	P11667 escherichia
4	319	26.8	201	1 Y488 MYCTU	Q11154 mycobacteri
5	318	26.7	199	1 Y488 MYCTU	Q10871 mycobacteri
6	310	26.0	206	1 YGGA_AERSA	P70775 aeromonas s
7	254	21.3	225	1 YGGA_AERHY	P52047 aeromonas h
8	105	8.8	206	1 RHFC_ECOLI	P27846 escherichia
9	101	8.5	212	1 YEAS_ECOLI	P76249 escherichia
10	93.5	7.9	452	1 PTCC_ECOLI	P17334 escherichia
11	92.5	7.8	465	1 Y093_RHIME	O87394 rhizobium m
12	92	7.7	223	1 YAHN_ECOLI	P75693 escherichia
13	91	7.6	206	1 RHFC_SALTY	Q916n7 salmonella
14	89	7.5	944	1 CHS2_NEUCR	P30589 neurospora
15	88.5	7.4	3137	1 CA36_CHICK	P15989 gallus gall
16	88	7.4	206	1 RHFC_SALTY	Q8z3b3 salmonella
17	87.5	7.3	443	1 DCUA_HELPY	O25425 helicobacte
18	87.5	7.3	452	1 DNAA_STRMU	Q8dwn9 streptococc
19	87	7.3	295	1 YBIF_SALTY	P74867 salmonella
20	86.5	7.3	195	1 YBIF_ECOLI	P38101 escherichia
21	86	7.2	295	1 YBIF_ECOLI	P36545 escherichia
22	86	7.2	351	1 MRAY_METHA	O26830 methanobact
23	85	7.1	302	1 Y091_METHA	Q57556 methanococc
24	84.5	7.1	253	1 CYSZ_ECOLI6	Q8fbt9 escherichia
25	84.5	7.1	443	1 DCUA_HELPY	Q9z1c0 helicobacte
26	84	7.1	482	1 YFIG_BACSU	P54723 bacillus su
27	84	7.1	685	1 FHUB_SALTY	O87656 salmonella
28	83.5	7.0	725	1 HPFA_RHOFA	Q8ky01 rhodospheudo
29	83	7.0	216	1 YBIF_PSEAE	P38102 pseudomonas
30	83	7.0	726	1 HPFA_THEMEA	O9s5x0 thermotoga
31	82.5	6.9	3491	1 ERV1_SACER	Q03131 saccharopol
32	81.5	6.8	253	1 CYSZ_SALTY	P8z4w3 salmonella
33	81.5	6.8	742	1 DHET_ACEAC	P18278 acetobacter

Query Match 98.9% Score 1178; DB 1; Length 233;

34	81	6.8	622	1 KUP_ECO57	Q8xw9 escherichia
35	81	6.8	622	1 KUP_ECOLI	P30016 escherichia
36	81	6.8	676	1 COMF_RHIME	P45404 rhizobium m
37	80.5	6.8	253	1 CYSZ_ECOLI	P12610 escherichia
38	80.5	6.8	253	1 CYSZ_SALTY	P12673 salmonella
39	80.5	6.8	456	1 AROP_ECOLI6	Q8f149 escherichia
40	80.5	6.8	532	1 VAT_HUMAN	Q16572 homo sapien
41	80	6.7	206	1 RHTE_ECOLI	P27847 escherichia
42	80	6.7	394	1 YGAY_ECO57	Q8x4v6 escherichia
43	80	6.7	461	1 CSBC_BACSU	P46333 bacillus su
44	80	6.7	530	1 VAT_MOUSE	O35304 mus musculu
45	79.5	6.7	215	1 ORN_MYCLE	O07708 mycobacteri

ALIGNMENTS

RESULT 1	LYSE CORGL	STANDARD;	PRT;	233 AA.
ID	AC P94633;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Lysine exporter protein.			
GN	LYSE OR CG11262.			
OS	Corynebacterium glutamicum (Brevibacterium flavum).			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.			
OX	NCBI_TaxID=1718;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND FUNCTION.			
RC	STRAIN=RL27;			
RX	MEDLINE=97126810; PubMed=8971704;			
RA	Vrljic M.M., Sahm H., Eggeling L.;			
RT	"A new type of transporter with a new type of cellular function: L-			
RT	lysine export from Corynebacterium glutamicum."			
RL	Mol. Microbiol. 22:815-826(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;			
RA	Nakagawa S.;			
RT	"Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."			
RT	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: INVOLVED IN THE EFFLUX OF EXCESS OF L-LYSINE. THIS IS			
CC	NECESSARY TO CONTROL THE INTRACELLULAR L-LYSINE LEVEL.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.			
CC	-!- SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X96471; CAA65324.2; ..			
DR	EMBL; AP005277; BAB98655.1; ALT INIT.			
DR	InterPro; IPR004777; Lys exporter.			
DR	InterPro; IPR001123; LysE.			
DR	Pfam; PF01810; LysE; 1.			
DR	TIGRFAMs; TIGR00948; 2a75; 1.			
KW	Transport; Transmembrane; Inner membrane; Complete proteome.			
FT	TRANSMEM 3 23 POTENTIAL.			
FT	TRANSMEM 35 55 POTENTIAL.			
FT	TRANSMEM 66 86 POTENTIAL.			
FT	TRANSMEM 144 164 POTENTIAL.			
FT	TRANSMEM 177 197 POTENTIAL.			
FT	TRANSMEM 213 233 POTENTIAL.			
SQ	SEQUENCE 233 AA; 25082 MW; F5FD9B1ACAD1D13 CRC64;			

Best Local Similarity 100.0%; Pred. No. 7.1e-95;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEIPIITGLGASLLLSIGFQNVYIKQIKREGLIAVLVCLISDVFLFIAGTGLVDLL 63
Db 1 MEIPIITGLGASLLLSIGFQNVYIKQIKREGLIAVLVCLISDVFLFIAGTGLVDLL 60

QY 64 SNAAPITVLDIMRWGIAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPLGSAVA 123
Db 61 SNAAPITVLDIMRWGIAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPLGSAVA 120

QY 124 TDTNRNVRVSVKQVQVWVKPMLMAIIVLTWLNPNAYLDAFAVFIIGVGAQYGTGRWIFA 183
Db 121 TDTNRNVRVSVKQVQVWVKPMLMAIIVLTWLNPNAYLDAFAVFIIGVGAQYGTGRWIFA 180

QY 184 AGAFAASLIWFLVFGGAALSRLSPSSPKVWRWVNVVAVMTALAIAKMLMG 236
Db 181 AGAFAASLIWFLVFGGAALSRLSPSSPKVWRWVNVVAVMTALAIAKMLMG 233

RESULT 2
LYSE COREF
ID LYSE COREF STANDARD; PRT; 228 AA.
AC Q8RQ4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lysine exporter protein.
GN LYSE OR CBI357.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579(2003).
RN [2]
RP "FUNCTION: INVOLVED IN THE EFFLUX OF EXCESS OF L-LYSINE. THIS IS
CC NECESSARY TO CONTROL THE INTRACELLULAR L-LYSINE LEVEL (By
CC similarity).
CC (By similarity).
CC -!- SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB088333; BAB88827.1; ALT_INIT.
CC EMBL; AF005218; BAC18167.1; ALT_INIT.
CC InterPro; IPR01123; Lyse.
CC Pfam; PF01810; Lyse; 1.
CC Transport; Transmembrane; Inner membrane; Complete proteome.
CC
CC TRANSMEM 3 23 POTENTIAL.
CC FT TRANSMEM 37 57 POTENTIAL.
CC FT TRANSMEM 66 86 POTENTIAL.

FT TRANSMEM 139 159 POTENTIAL.
FT TRANSMEM 172 192 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
SQ SEQUENCE 228 AA; 24374 MW; 68FE379A830EDA7 CRC64;
Query Match 72.6%; Score 864.5; DB 1; Length 228;
Best Local Similarity 71.2%; Pred. No. 9.1e-68;
Matches 166; Conservative 30; Mismatches 32; Indels 5; Gaps 2;

QY 4 MEIPIITGLGASLLLSIGFQNVYIKQIKREGLIAVLVCLISDVFLFIAGTGLVDLL 63
Db 1 MEIPIITGLGASLLLSIGFQNVYIKQIKREGLIAVLVCLISDVFLFIAGTGLVDLL 60

QY 64 SNAAPITVLDIMRWGIAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPLGSAVA 123
Db 61 SNAAPITVLDIMRWGIAIYLLWFAVMAAKDAMTKVEAPQIIEETPTVDDTPLGSAVA 115

QY 124 TDTNRNVRVSVKQVQVWVKPMLMAIIVLTWLNPNAYLDAFAVFIIGVGAQYGTGRWIFA 183
Db 116 VTTKQRPRLRITSGTRQVWVRPMLMAIIVLTWLNPNAYLDAFAVFIIGVGAQYGTGRWIFA 175

QY 184 AGAFAASLIWFLVFGGAALSRLSPSSPKVWRWVNVVAVMTALAIAKMLMG 236
Db 176 AGAFAASLIWFLVFGGAALSRLSPSSPKVWRWVNVVAVMTALAIAKMLMG 228

RESULT 3
YGGG ECOLI
ID YGGG ECOLI STANDARD; PRT; 211 AA.
AC P11667;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yggA.
GN YGGG OR B2923.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
RN [2]
RP SEQUENCE OF 15-211 FROM N.A.
RC STRAIN=K12 / CS520;
RX MEDLINE=89313302; PubMed=2546007;
RA Alefounder P.R., Baldwin S.A., Perham S.A., Short N.J.;
RT "Identification, molecular cloning and sequence analysis of a gene
RT cluster encoding the class II fructose 1,6-bisphosphate aldolase, 3-
RT phosphate dehydrogenase and a putative second glyceraldehyde 3-
RT Mol. Microbiol. 3:723-732 (1989).
RL -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U28377; AAA69090.1; -.
CC EMBL; AB000375; AAC75960.1; -.
CC EMBL; X14436; CAA32607.1; -.
CC
CC DR EMBL; U28377; AAA69090.1; -.
CC DR EMBL; AB000375; AAC75960.1; -.
CC DR EMBL; X14436; CAA32607.1; -.

```

DR PIR; B65077; QORCEA.
DR EcoGene; EG1159; Y9GA.
DR InterPro; IPR004777; Lys exporter.
DR InterPro; IPR001123; LysE.
DR Pfam; PF01810; LysE; 1.
DR TIGRFAMs; TIGR00948; 2a75; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 111 131 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
SQ SEQUENCE 211 AA; 23175 MW; 2DAFE27B6A9B5822 CRC64;

Query Match 29.4%; Score 350; DB 1; Length 211;
Best Local Similarity 35.7%; Pred. No. 2.7e-23;
Matches 81; Conservative 37; Mismatches 75; Indels 34; Gaps 4;

QY 7 FITGLLGLSLLSTGPQNVLVKQIKREGILAVLVCLISDVFLFIAGTLGVDLLSNA 66
DB 5 YFQGLAAGAAVILPLGPQNAFVMQGRIRQYHIMALLCAISDLVLCAGIFGGSALLMQ 64
QY 67 APVLIDIMRWGGTAYLLFVFAVAAKADMTNKVEAPQIIETEPTVPDTPLGSAVATDT 126
DB 65 SPWLLALVWGGVAFLLWYGGAFKTMSSNIEL----- 98
QY 127 RNRVREVSVDKQVWVKMELMAVLFWLNPNAVLDVFFVGGVGAQYG-DTGRWIFPAAG 185
DB 99 -----ASAEVWKQGRW-KIATMLAVTLNPHVYLDTEVLGSLGGQLDVEPKRW-FALG 151
QY 186 AFASLWFLVFGFGAALSRLSSPKVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVW 232
DB 152 TISASFLWFFGLALLAALAPRLTAKAQRILNVWGVWVWVWVWVWVWVWVWVWVWVWVW 198

RESULT 4
Y488 MYCTU STANDARD; PRT; 201 AA.
AC Q11154;
DT 01-OCT-1996 (Rel. 34, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Hypothetical protein Rv0488/MT0507/Mb0498.
GN Rv0488 OR MT0507 OR MTCY20G9.14 OR Mb0498.
OS Mycobacterium tuberculosis, and
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

```

```

RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE LYSY/YGGA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z77162; CAB00949.1; -.
CC EMBL; AE006952; RAK44730.1; -.
CC EMBL; BX248335; CAD93361.1; -.
CC PIR; C70744; C70744.
CC TIGR; MT0507; -.
CC TubercuList; Rv0488; -.
CC InterPro; IPR004777; Lys exporter.
CC InterPro; IPR001123; LysE.
CC Pfam; PF01810; LysE; 1.
CC TIGRFAMs; TIGR00948; 2a75; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 25 45 POTENTIAL.
CC FT TRANSMEM 57 77 POTENTIAL.
CC FT TRANSMEM 104 124 POTENTIAL.
CC FT TRANSMEM 133 153 POTENTIAL.
CC FT TRANSMEM 169 189 POTENTIAL.
CC SQ SEQUENCE 201 AA; 20951 MW; E198975DF088B6E4 CRC64;

Query Match 26.8%; Score 319; DB 1; Length 201;
Best Local Similarity 32.4%; Pred. No. 1.2e-20;
Matches 71; Conservative 50; Mismatches 64; Indels 34; Gaps 4;

QY 16 SLLSISGPQNVLVKQIKREGILAVLVCLISDVFLFIAGTLGVDLLSNAFVLDIMR 75
DB 3 TLKVAIGPQNAFVLRQIRREYVLVIVALGCIADGALIAAGVGGFAALIIHAHPNMTLVAR 62
QY 76 WGGIAYLLWFAVMAKADMTNKVEAPQIIETEPTVPDTPLGSAVATDTNRVREVS 135
DB 63 FGGAFLGIYALLAARNW-----RPSGLVPSES--GPAAL----- 96
QY 136 VDKQRVWVKMELMAVLFWLNPNAVLDVFFVGGVGAQYGTGRWIFPAAGFAASLIWFP 195
DB 97 -----IGVQMCVLTVPFLNPHVYLDTEVLGSLGGQLDVEPKRW-FALG 148
QY 196 LVGFGAALSRLSSPKVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVW 234
DB 149 VLGSAGRLQPPFATPAARVILDLAVMTVIGVAVVVLV 187

RESULT 5
YJ86 MYCTU STANDARD; PRT; 199 AA.
AC Q10871;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-2003 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Rv1986/MT2040/Mb2008.
GN Rv1986 OR MT2040 OR MTCY39.33C OR Mb2008.

```


Best Local Similarity 31.1%; Pred. No. 1.5e-20;
Matches 70; Conservative 44; Mismatches 77; Indels 34; Gaps 3;

Qy 8 ITGILLGASLLISGPONVLYIKQIKREGLIAVLAVCLISDVFLFTAGTILGVDLISNAA 67
:
Db 6 VVGFLACFTLIAAIGAQNAFVLRQIQIQRHVLPVALCTVSIVLIAAGIAGFGALICAH 65
:
Qy 68 PIVLIDIMRWGGIAYLVLFWFAAKADMTNKVEAPQIIERTPTPDDTPLGGSVAVDTR 127
:
Db 66 PRALNVKKGGAFLGYGLLAAREWPVALIP-----SGATPVR 106
:
Qy 128 NRVREVSVDRQVVVKPMMLAIIVLTWLNPNAYLDFAVFVIGVGAQYGDTRWIFAACAF 187
:
Db 107 -----LAELVTVCRAFTFLNPHEVLDTWLVLGALANESDQ-RWIFLGAV 151
:
Qy 198 AASLIPLFVLGVGAALSRPLSPKWVINVVAVVMATAIKL 232
:
Db 152 TASAVWFATLGFAGRLGRGLTFNGSWRIIDGLIANVVALGISL 196
: :

RESULT 6

YGGA AERSA STANDARD; PRT; 206 AA.

ID YGGA AERSA STANDARD; PRT; 206 AA.
AC F70775;
CD 15-JUL-1998 (Rel. 36, Created)
DDT 15-JUL-1998 (Rel. 36, Last sequence update)
EED 28-FEB-2003 (Rel. 41, Last annotation update)
FFT Hypothetical protein yggA.
GGG YGGA.
HHH Aeromonas salmonicida.
III Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;
KKK Aeromonadaceae; Aeromonas.
LLL NCBI_TaxID=645;
MMM SEQUENCE FROM N.A.
NNN STRAIN-NCIMB 1102;
OOO MEDLINE=97431471; PubMed=9286976;
PPR Swift S., Kariyeh A.V., Fish L., Durant E.L., Winson M.K.,
RRR Chhabra S.R., Williams P., Macintyre S., Stewart G.S.A.B.;
SSS "Quorum sensing in Aeromonas hydrophila and Aeromonas salmonicida:
TTT identification of the LuxRI homologs AhvRI and AsaRI and their
UUU cognate N-acylhomoserine lactone signal molecules.";
VVV J. Bacteriol. 179:5271-5281(1997).
WWW - SUBCELLULAR LOCATION: Integral membrane protein (Probable).
XXX - SIMILARITY: BELONGS TO THE LYSE/YGGA FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).

EMBL: U65741; AAB70019.1; ALT INIT.
InterPro: IPR004777; Lys exporter.
InterPro: IPR001123; LytE.
Pfam: PF01810; Lyse; 1.
TIGRfam: TIGR00948; za75; 1.
Hypothetical protein; Transmembrane.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 116 136 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
SEQUENCE 206 AA; 21505 MW; D1C2CA92CDA0179A CRC64;

Query Match 26.0%; Score 310; DB 1; Length 206;
Best Local Similarity 31.9%; Pred. No. 7.6e-20;
Matches 72; Conservative 43; Mismatches 77; Indels 34; Gaps 3;

Ov 8 ITGILLGASLLISGPONVLYIKQIKREGLIAVLAVCLISDVFLFTAGTILGVDLISNAA 67

```

Db      6  LQFTGLGAMIPIGAQNAFVLSRGIRHNHLLAATLCCCLDLILIGVFGGANLLAAS 65
QY      68  PIVLDIMRWGGIAYLLWFVFAAKADMTNKVEAPQIIEETPTVPDDTPPGSSAVATDTR 127
Db      66  PIGLALLTWGGVFLCNGFIRSLSAQGGQ-----GAALADGPR 104
QY      128  NRVREVSVDKQVRWKPKML-NAIVLTWLNPNAYLDADFVIGGCAQYGDGTGWIFAAGA 186
Db      105  -----LMGVKSVLMTLGVTLNPHVLTDTMLLGSFGSGFAPLRPAAGA 152
QY      187  FAASLTWFLPFGGAALSRPLSSPKWRNINNVAVVMTALAIAKL 232
Db      153  MLASLVWFVSLAFGAALSPWLAGRGVQAIDTIVGLMGLAQQL 198

RESULT 7
YCGA AERHY STANDARD; PRT; 225 AA.
AC P52047;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Hypothetical 24.5 kDa protein in ahyR-cdpD intergenic region.
OS Aeromonas hydrophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;
OC Aeromonadaceae; Aeromonas.
OX NCBI_TaxID=644;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97431471; PubMed=9286976;
RA Swift S., Karlyshev A.V., Fish L., Durant E.L., Winson M.K.,
RA Chhabra S.R., Williams P., Macintyre S., Stewart G.S.A.B.;
RT "Quorum sensing in Aeromonas hydrophila and Aeromonas salmonicida:
RT identification of the LuxR homologs AhyRI and AhyRI and their
RT cognate N-acylhomoserine lactone signal molecules";
RL J. Bacteriol. 179:5271-5281(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE LYSE/YCGA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X89469; CAA61655.1; ALT INIT.
CC InterPro; IPR004777; Lys exporter.
CC InterPro; IPR001123; LyseE.
CC Pfam; PF01810; LyseE; 1.
CC TIGRFAMs; TIGR00948; 2a75; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 116 136 POTENTIAL.
FT TRANSMEM 150 170 POTENTIAL.
SQ SEQUENCE 225 AA; 24482 MW; 172DB104473B0809 CRC64;

Query Match 21.3%; Score 254; DB 1; Length 225;
Best Local Similarity 26.9%; Pred. No. 5,8e-15;
Matches 61; Conservative 38; Mismatches 74; Indels 54; Gaps 4;

QY      8  ITGLLGASLLSISGPQNVLIQKQIKREGIAYLVCLISDVFLFIAGTLGVLLSNRA 67
Db      6  LQFTGLGAMIPIGAQNAFVLSRGIRHNHLLAATLCCCLDLILIGVFGGANLLAAS 65
QY      68  PIVLDIMRWGGIAYLLWFVFAAKADMTNKVEAPQIIEETPTVPDDTPPGSSAVA 123
Db      66  PIGLALLTWGGVFLCNGFIRSLSAQGGQ-----MGVKSYS- 111

```

```

QY      124  TDTNRVRVSVSKQVRWKPKMLNAIVLTWLNPNAYLDADFVIGGCAQYGDGTGWIFA 183
Db      112  -----LMTLGVTLNPHVLTDTMLLGSFGSGFAELASAF 149
QY      184  AGAFAASLTWFLPFGGAALSRPLSSPK-----VMRW 216
Db      150  AVAMLASLVWFVSLAFGAALSPWLAGRGVQAIDTIVGLMGLAQQL 196

RESULT 8
RHTC ECOLI STANDARD; PRT; 206 AA.
AC P27846;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Threonine efflux protein.
GN RHTC OR B3823 OR Z5344 OR ECS4753.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes.";
RL Science 257:771-778(1992).
RN [2]
RP REVISIONS.
RC STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [5]
RP SEQUENCE OF 1-107 FROM N.A.
RC STRAIN=K12;
RC MEDLINE=87115164; PubMed=3027506;
RA Irino N., Nakayama K., Nakayama H.;
RT "The reqQ gene of Escherichia coli K12: primary structure and
RT evidence for SOS regulation.";
RL Mol. Gen. Genet. 205:298-304(1986).

```

RESULT 9

QY 179 RWIFAA-----GAPAA-----SLWFLVGPAA 201
 DB 153 PFILANTLELVSFYLISLISGAFVQVIRTKKLAKVNSLGLMLFVFAA 205

RESULT 10
 PTCC_ECOLI STANDARD; PRT; 452 AA.
 AC P17334; P76212; P76907; P77332;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE PPS system, N,N'-diacetylchitobiose-specific IIC component (EIIIC-Chb)
 DE (N,N'-diacetylchitobiose-permease IIC component) (Phosphotransferase
 DE enzyme II, C component).
 GN CHBC OR CELB OR B1737.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12; PubMed=2179047;
 RA Parker L.L., Hall B.G.;
 RT "Characterization and nucleotide sequence of the cryptic cel operon
 RT of Escherichia coli K12.";
 RL Genetics 124:455-471 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sempel G., Seki Y., Sivasaundaram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377 (1996).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=91227627; PubMed=2092358;
 RA Reizer J., Reizer A., Sailer M.H. Jr.;
 RT "The cellobiose permease of Escherichia coli consists of three
 RT proteins and is homologous to the lactose permease of Staphylococcus
 RT aureus.";
 RL Res. Microbiol. 141:1061-1067 (1990).
 RN [5]
 RP IDENTIFICATION OF CHB OPERON.
 RX MEDLINE=98070757; PubMed=9405618;
 RA Keyhani N.O., Roseman S.;
 RT "Wild-type Escherichia coli grows on the chitin disaccharide,
 RT N,N'-diacetylchitobiose, by expressing the cel operon.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:14367-14371 (1997).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=20490752; PubMed=10913117;
 RA Keyhani N.O., Wang L.-X., Lee Y.C., Roseman S.;
 RT "The chitin disaccharide, N,N'-diacetylchitobiose, is catabolized by
 RT Escherichia coli and is transported/phosphorylated by the
 RT phosphoenolpyruvate:glycose phosphotransferase system.";

J. Biol. Chem. 275:33084-33090 (2000).
 CC -!- FUNCTION: This is a component of the phosphoenolpyruvate-dependent
 CC sugar phosphotransferase system (PTS), a major carbohydrate active
 CC -transport system. The IIC domains contain the sugar binding site
 CC and the transmembrane channel; the IIA domain contains the primary
 CC phosphorylation site (the donor is phospho-HPr); IIA transfers its
 CC phosphoryl group to the IIB domain which finally transfers it to
 CC the sugar.
 CC -!- PATHWAY: N,N'-diacetylchitobiose utilization.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -!- INDUCTION: By N,N'-diacetylchitobiose.
 CC -!- SIMILARITY: Contains 1 PTS EIIIC domain.
 CC -!- CAUTION: Was originally (Ref.4) characterized as part of a cryptic
 CC cel operon for a cellobiose degradation system. The Cel₁ phenotype
 CC is due to mutations making expression chitobiose-independent and
 CC altering the substrate specificity.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X52890; CAA37070.1; -;
 CC EMBL; A5000269; AAC74807.1; -;
 CC EMBL; D90816; BAA15518.1; -;
 CC EMBL; D90817; BAA15526.1; -;
 CC EMBL; D90818; BAA15532.1; -;
 CC PIR; A64933; A64933.
 CC PIR; S10871; S10871.
 CC EcoGene; EGI0141; chbC.
 CC InterPro; IPR004796; Cello_Pts_IIC.
 CC InterPro; IPR003352; Ptrans_EIIC.
 CC InterPro; IPR004501; Pts_lac.
 CC Pfam; PF02378; Pts_EIIC; 1.
 CC TIGRFAMs; TIGR00359; Cello_Pts_IIC; 1.
 CC TIGRFAMs; TIGR00410; lacB; 1.
 CC Phosphotransferase system; Sugar transport; Transmembrane;
 CC Inner membrane; Complete proteome.
 CC TRANSMEM 35 55 POTENTIAL.
 CC TRANSMEM 86 106 POTENTIAL.
 CC TRANSMEM 114 134 POTENTIAL.
 CC TRANSMEM 147 167 POTENTIAL.
 CC TRANSMEM 187 207 POTENTIAL.
 CC TRANSMEM 229 249 POTENTIAL.
 CC TRANSMEM 300 320 POTENTIAL.
 CC TRANSMEM 344 364 POTENTIAL.
 CC TRANSMEM 367 387 POTENTIAL.
 CC TRANSMEM 409 429 POTENTIAL.
 CC CONFLICT 190 205 LIPGFIILSVNGIAW -> FNSRLVYSFRDGYCL (IN
 CC REF. 1).
 CC CONFLICT 233 235 VGM -> WL (IN REF. 1).
 CC CONFLICT 240 265 FVPLLVFGHIALTALDNGIMTP -> LSTALVIRIHA
 CC ACADLTGQHYDA (IN REF. 1).
 CC CONFLICT 361 366 FVLVQP -> LYVYNR (IN REF. 1).
 CC CONFLICT 405 452 SVAALLVALFNIGIATLYLPVVVANKAQNADKESSED
 CC IANALKF -> TSPHCWSSHTLASQR (IN REF. 1).
 CC SEQUENCE 452 AA; 48332 MW; EC3CBFEED6231068 CRC64;
 Query Match 7.9%; Score 93.5; DB 1; Length 452;
 Best local Similarity 19.4%; Pred. No. 0.9;
 Matches 52; Conservative 44; Mismatches 95; Indels 77; Gaps 14;
 QY 8 ITGLLIG--ASLLLSIGQNVLV-----IKQGIKR--EGLIALLVCLISDYFLIAGT 57
 DB 150 ISGIIGLWAEFTFIVRNWVIXLPSVPASVSFSLIPGFIILSVNGIAWNT 209
 QY 58 LG-----VLLSNRAPIVLDMRWGGIAY--LLWFAVW---AAKDAMTKVEAPQIIE 105
 DB 210 WGTNFHQIMDTITPLASIGSVGWAVFVPLLVFFGHGHALTALDNGIMTPWALE 269

QY 106 ETEPTVDDTGLGSAVAT-DTNRNRVVEVSVDKQ-RVMVKPMLMAIVLTLNPNAYLDA 163
 Db 270 N-----IATYQVGSVEALAAAGKTHIWAKM-----LDS 300
 QY 164 FVFIGGVG-----AQYDGTGRWIFAFAGAFASLIWFFPLVFGGAAALSRPL 208
 Db 301 FIFLGSGATLGLILAIPIASRRADYQVAKLALPSGIFQ---INEPIL-FGLPIIMNFV 356
 QY 209 SSFQVRWVNVVAVMTALAIKLMG 236
 Db 357 MFIFP-----VLVQPIILAAITLAAYTMG 379
 RESULT 11
 Y093_RHIME
 ID Y093_RHIME STANDARD; PRT; 465 AA.
 AC 087394;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical transport protein R00093.
 GN R00093 OR SMC02616.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA Powers E.L.; Vuyyuru V., Kahn M.L.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capella D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetlelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 RT Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.";
 RC Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 RL -!- FUNCTION: Probable amino-acid or metabolite transport protein.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the amino acid permease family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF055582; AAC62224.1; -
 CC EMBL; AL591782; CAC41480.1; -
 CC InterPro; IPR002293; AA/rel permease1.
 CC InterPro; IPR004840; AAC permease.
 CC InterPro; IPR004841; Permease region.
 CC Pfam; PF00324; aa_permeases; 1.
 CC PROSITE; PS00218; AMINO-ACID PERMEASE 1; FALSE NEG.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSNMEN 19 POTENTIAL.
 FT TRANSNMEN 50 70
 FT TRANSNMEN 91 111
 FT TRANSNMEN 140 160
 FT TRANSNMEN 164 184
 FT TRANSNMEN 201 221
 FT TRANSNMEN 244 264
 FT TRANSNMEN 288 308
 FT TRANSNMEN 342 362
 FT TRANSNMEN 363 383

FT TRANSNMEN 403 423 POTENTIAL.
 SQ SEQUENCE 465 AA; 50783 MW; 0375E164F737AA0A CRC64;
 Query Match 7.8%; Score 92.5; DB 1; Length 465;
 Best Local Similarity 20.0%; Pred. No. 1.1;
 Matches 47; Conservative 38; Mismatches 83; Indels 67; Gaps 7;
 QY 17 LLISIGPQNVLVIKQI-----KREGLIAVLVCLISDFLFTAGTGLGVDLL 63
 Db 16 LLRVLGPAHVALGVIVLVEYMGWFSVGKGIAGLMACWAGLLTYTCVAMIDSEVT 75
 QY 64 SNRAP-----IVLDIMRWG-GIAYLLWFAVMAAKDMNKVEAPQIIEETPTVP 112
 Db 76 STVAAGGQYQAQKHIVGPLMAFNVLFLVWAYTLEAANAIT----- 118
 QY 113 DDTPLGSAVATDTNRNRVVEVSVDKQ-RVMVKPMLMAIVLTLNPNAYLDAFVFIGVGVA 172
 Db 119 -----VGFLDVTAGVGGQTGLNQPPIV---LAIPLFLANLNRGLVATLTF----- 162
 QY 173 QYDGTGRWIFAFAGAFASLIWFFPLVFGGAAAL-----SRPLSSPKVWRWNVVAV 223
 Db 163 -----NLVITAIPLAIVLAVFVSQFGASAVPLDFAITSDPLPYGWWGVASLI 211
 RESULT 12
 YAHN_ECOLI
 ID YAHN_ECOLI STANDARD; PRT; 223 AA.
 AC P75693; P71307;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yahn.
 GN YAHN OR B0328.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 [2]
 RP SEQUENCE FROM N.A.
 RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
 RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H.,
 RA Lin D., Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the rht family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE000140; AAC73431.1; -
 CC EMBL; U73857; AAB18053.1; -
 CC PIR; H64759; H64759.
 CC Ecogene; EGI3598; yahn.
 CC InterPro; IPR004778; Homoser_Thr_eff.
 CC InterPro; IPR001123; Lyse.
 CC Pfam; PF01810; Lyse; 1.
 CC TIGRFAMs; TIGR00949; 2A76; 1.
 KW Hypothetical protein; Transmembrane; Complete proteome.

```

FT TRANSMEM      1  21          POTENTIAL..
FT TRANSMEM     44   64          POTENTIAL..
FT TRANSMEM     67   87          POTENTIAL..
FT TRANSMEM    150  173          POTENTIAL..
SQ SEQUENCE     206 AA; 22498 MW; 080477853FC2733F CRC64;

Query Match              7.6%; Score 91; DB 1; Length 206;
Best Local Similarity   20.3%; Pred. No. 0.66;
Matches 50; Conservative 34; Mismatches 106; Indels 56; Gaps 8

QY 3 IMEITGLLGASLLSTGPNVLVIKQIKREGILAVLLVCLISDVFLFAGT--LGV 60
Db 1 MMMLRFVTVMVHVHVALMSGPDPFFVSQTAVSRKEAMGVLTGTGGVWVWAGVALLGL 60
QY 61 DLLSNAAPLVLDIMRWGGGIAYLLMFAYMAKAQAMTNK---VEAQIIBETFTVPDDTPL 117
Db 61 HLIIIEKMAWLHTTIIVGGGGLYLWNMGVQLRGALKKKQAAAASSPHI----- 106
QY 118 GGSNAVTDTRNRVRVESVDQRVVVQPMLEMAVLITWLNPNAYLDFAVFIGG----VGA 172
Db 107 ---ELAQSGHS-----FLKGILTLSNPK---AIIYFGSVFSLEFVG 142
QY 173 QYGDTGRMIFAAAGAFA---ASLIWFPLVCGAALSRPLSSPKWRNVINVVAVVMTAL 228
Db 143 NVGAARW---GIPALITLETLANFTVVASULFALPKWRGRYORLAKWIDGFAGALPAGE 198
QY 229 ATKML 234
Db 199 GIHLII 204
```

```

RESULT 14
CHS2_NEUCR
ID CHS2_NEUCR STANDARD; PRT; 944 AA.
AC P30589;
DT DT 01-APR-1993 (Rel. 25, Created)
DT DT 01-APR-1996 (Rel. 33, Last sequence update)
DT DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Chitin synthase 2 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
DE transferase 2).
DE CHS-2.
GN CHS-2.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariaceae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RN [2]
RN [3]
RN [4]
RN [5]
RN [6]
RN [7]
RN [8]
RN [9]
RN [10]
RN [11]
RN [12]
RN [13]
RN [14]
RN [15]
RN [16]
RN [17]
RN [18]
RN [19]
RN [20]
RN [21]
RN [22]
RN [23]
RN [24]
RN [25]
RN [26]
RN [27]
RN [28]
RN [29]
RN [30]
RN [31]
RN [32]
RN [33]
RN [34]
RN [35]
RN [36]
RN [37]
RN [38]
RN [39]
RN [40]
RN [41]
RN [42]
RN [43]
RN [44]
RN [45]
RN [46]
RN [47]
RN [48]
RN [49]
RN [50]
RN [51]
RN [52]
RN [53]
RN [54]
RN [55]
RN [56]
RN [57]
RN [58]
RN [59]
RN [60]
RN [61]
RN [62]
RN [63]
RN [64]
RN [65]
RN [66]
RN [67]
RN [68]
RN [69]
RN [70]
RN [71]
RN [72]
RN [73]
RN [74]
RN [75]
RN [76]
RN [77]
RN [78]
RN [79]
RN [80]
RN [81]
RN [82]
RN [83]
RN [84]
RN [85]
RN [86]
RN [87]
RN [88]
RN [89]
RN [90]
RN [91]
RN [92]
RN [93]
RN [94]
RN [95]
RN [96]
RN [97]
RN [98]
RN [99]
RN [100]
RN [101]
RN [102]
RN [103]
RN [104]
RN [105]
RN [106]
RN [107]
RN [108]
RN [109]
RN [110]
RN [111]
RN [112]
RN [113]
RN [114]
RN [115]
RN [116]
RN [117]
RN [118]
RN [119]
RN [120]
RN [121]
RN [122]
RN [123]
RN [124]
RN [125]
RN [126]
RN [127]
RN [128]
RN [129]
RN [130]
RN [131]
RN [132]
RN [133]
RN [134]
RN [135]
RN [136]
RN [137]
RN [138]
RN [139]
RN [140]
RN [141]
RN [142]
RN [143]
RN [144]
RN [145]
RN [146]
RN [147]
RN [148]
RN [149]
RN [150]
RN [151]
RN [152]
RN [153]
RN [154]
RN [155]
RN [156]
RN [157]
RN [158]
RN [159]
RN [160]
RN [161]
RN [162]
RN [163]
RN [164]
RN [165]
RN [166]
RN [167]
RN [168]
RN [169]
RN [170]
RN [171]
RN [172]
RN [173]
RN [174]
RN [175]
RN [176]
RN [177]
RN [178]
RN [179]
RN [180]
RN [181]
RN [182]
RN [183]
RN [184]
RN [185]
RN [186]
RN [187]
RN [188]
RN [189]
RN [190]
RN [191]
RN [192]
RN [193]
RN [194]
RN [195]
RN [196]
RN [197]
RN [198]
RN [199]
RN [200]
RN [201]
RN [202]
RN [203]
RN [204]
RN [205]
RN [206]
RN [207]
RN [208]
RN [209]
RN [210]
RN [211]
RN [212]
RN [213]
RN [214]
RN [215]
RN [216]
RN [217]
RN [218]
RN [219]
RN [220]
RN [221]
RN [222]
RN [223]
RN [224]
RN [225]
RN [226]
RN [227]
RN [228]
RN [229]
RN [230]
RN [231]
RN [232]
RN [233]
RN [234]
RN [235]
RN [236]
RN [237]
RN [238]
RN [239]
RN [240]
RN [241]
RN [242]
RN [243]
RN [244]
RN [245]
RN [246]
RN [247]
RN [248]
RN [249]
RN [250]
RN [251]
RN [252]
RN [253]
RN [254]
RN [255]
RN [256]
RN [257]
RN [258]
RN [259]
RN [260]
RN [261]
RN [262]
RN [263]
RN [264]
RN [265]
RN [266]
RN [267]
RN [268]
RN [269]
RN [270]
RN [271]
RN [272]
RN [273]
RN [274]
RN [275]
RN [276]
RN [277]
RN [278]
RN [279]
RN [280]
RN [281]
RN [282]
RN [283]
RN [284]
RN [285]
RN [286]
RN [287]
RN [288]
RN [289]
RN [290]
RN [291]
RN [292]
RN [293]
RN [294]
RN [295]
RN [296]
RN [297]
RN [298]
RN [299]
RN [300]
RN [301]
RN [302]
RN [303]
RN [304]
RN [305]
RN [306]
RN [307]
RN [308]
RN [309]
RN [310]
RN [311]
RN [312]
RN [313]
RN [314]
RN [315]
RN [316]
RN [317]
RN [318]
RN [319]
RN [320]
RN [321]
RN [322]
RN [323]
RN [324]
RN [325]
RN [326]
RN [327]
RN [328]
RN [329]
RN [330]
RN [331]
RN [332]
RN [333]
RN [334]
RN [335]
RN [336]
RN [337]
RN [338]
RN [339]
RN [340]
RN [341]
RN [342]
RN [343]
RN [344]
RN [345]
RN [346]
RN [347]
RN [348]
RN [349]
RN [350]
RN [351]
RN [352]
RN [353]
RN [354]
RN [355]
RN [356]
RN [357]
RN [358]
RN [359]
RN [360]
RN [361]
RN [362]
RN [363]
RN [364]
RN [365]
RN [366]
RN [367]
RN [368]
RN [369]
RN [370]
RN [371]
RN [372]
RN [373]
RN [374]
RN [375]
RN [376]
RN [377]
RN [378]
RN [379]
RN [380]
RN [381]
RN [382]
RN [383]
RN [384]
RN [385]
RN [386]
RN [387]
RN [388]
RN [389]
RN [390]
RN [391]
RN [392]
RN [393]
RN [394]
RN [395]
RN [396]
RN [397]
RN [398]
RN [399]
RN [400]
RN [401]
RN [402]
RN [403]
RN [404]
RN [405]
RN [406]
RN [407]
RN [408]
RN [409]
RN [410]
RN [411]
RN [412]
RN [413]
RN [414]
RN [415]
RN [416]
RN [417]
RN [418]
RN [419]
RN [420]
RN [421]
RN [422]
RN [423]
RN [424]
RN [425]
RN [426]
RN [427]
RN [428]
RN [429]
RN [430]
RN [431]
RN [432]
RN [433]
RN [434]
RN [435]
RN [436]
RN [437]
RN [438]
RN [439]
RN [440]
RN [441]
RN [442]
RN [443]
RN [444]
RN [445]
RN [446]
RN [447]
RN [448]
RN [449]
RN [450]
RN [451]
RN [452]
RN [453]
RN [454]
RN [455]
RN [456]
RN [457]
RN [458]
RN [459]
RN [460]
RN [461]
RN [462]
RN [463]
RN [464]
RN [465]
RN [466]
RN [467]
RN [468]
RN [469]
RN [470]
RN [471]
RN [472]
RN [473]
RN [474]
RN [475]
RN [476]
RN [477]
RN [478]
RN [479]
RN [480]
RN [481]
RN [482]
RN [483]
RN [484]
RN [485]
RN [486]
RN [487]
RN [488]
RN [489]
RN [490]
RN [491]
RN [492]
RN [493]
RN [494]
RN [495]
RN [496]
RN [497]
RN [498]
RN [499]
RN [500]
RN [501]
RN [502]
RN [503]
RN [504]
RN [505]
RN [506]
RN [507]
RN [508]
RN [509]
RN [510]
RN [511]
RN [512]
RN [513]
RN [514]
RN [515]
RN [516]
RN [517]
RN [518]
RN [519]
RN [520]
RN [521]
RN [522]
RN [523]
RN [524]
RN [525]
RN [526]
RN [527]
RN [528]
RN [529]
RN [530]
RN [531]
RN [532]
RN [533]
RN [534]
RN [535]
RN [536]
RN [537]
RN [538]
RN [539]
RN [540]
RN [541]
RN [542]
RN [543]
RN [544]
RN [545]
RN [546]
RN [547]
RN [548]
RN [549]
RN [550]
RN [551]
RN [552]
RN [553]
RN [554]
RN [555]
RN [556]
RN [557]
RN [558]
RN [559]
RN [560]
RN [561]
RN [562]
RN [563]
RN [564]
RN [565]
RN [566]
RN
```

```

CC -----
DR EMBL; X77782; CAA54816.1; -
DR EMBL; M82951; AAA33592.1; -
DR FIR; B45189; B45189.
DR FIR; T47246; T47246.
DR InterPro; IPR004834; Chitin synth.
DR Pfam; PF01644; Chitin synth; 1.
DR ProDom; PD002998; Chitin synth; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
FT TRANSMEM 597 617 POTENTIAL.
FT TRANSMEM 634 654 POTENTIAL.
FT TRANSMEM 669 689 POTENTIAL.
FT TRANSMEM 713 733 POTENTIAL.
FT TRANSMEM 873 893 POTENTIAL.
FT SEQUENCE 944 AA; 106816 MW; F70052AE0803060D CRC64;
SQ

Query Match 7.5%; Score 89; DB 1; Length 944;
Best Local Similarity 20.3%; Pred. No. 4.7;
Matches 60; Conservative 54; Mismatches 95; Indels 86; Gaps 14;

Qy 7 FITGLLGASLLISGPQNVIVIKQIKREGLIIVLVCLT-----SDVLFITAG----- 56
Db 641 YVCVLLICTQFISLGNR-----POGAKMYLASMIIYAVIMVYTTFTFIVVRQIQPS 695
Qy 57 -----TLGVLLSNAAPVILDMRWGGIAYLLMFVAAAKADMTNKVE-----APQI 103
Db 696 QKSDDKPDLNGLNVFTN---LIVSVASTLGLYFVMSFLYLDPMWMTISAQYFVLPSY 752
Qy 104 I-----ETETVPDD---TPLGSAVATDT----- 126
Db 753 ICTLIQIYAFCHDVITWGTGKGNVMTDLGGAIVKSGTVELEMPSPQLDIDSGYDECLRN 812
Qy 127 -RNRVVR---EYSDVK-ORWVVKEMLMALVLTWLNPNAYLDADFVIGGVAQYQDGTGRWI 181
Db 813 LRDKWVPCKSUSEQDQDYKSVRTYVWVSWVAVATLAM-----AVSAYGDSF--- 864
Qy 182 FAAGAFASLIIFPLVFGGAALSRPLSSPKVWRWVNVVAVVMTALAKLMLMG 236
Db 865 IGDNFYLRFLM-----AVALALFRLGSS-TTFAALNVLSVGLVGRVRLNMGK 914

RESULT 15
ID _CA36 CHICK STANDARD; PRT; 3137 AA.
AC P15989;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Collagen alpha 3(VI) chain precursor.
GN COL6A3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE OF 1-853 FROM N.A.
RC TISSUE=Heart;
RA MEDLINE=91035630; PubMed=1977751;
RA Doliana R., Bonaldo P., Colombatti A.;
RT "Multiple forms of chicken alpha 3(VI) collagen chain generated by
RT alternative splicing in type A repeated domains.";
RN J. Cell Biol. 111:2197-2205(1990).
RP [2]
RP SEQUENCE OF 224-2871 FROM N.A.
RA MEDLINE=90212613; PubMed=2322559;
RA Bonaldo P., Russo V., Buccioti F., Doliana R., Colombatti A.;
RT "Structural and functional features of the alpha 3 chain indicate a
RT bridging role for chicken collagen VI in connective tissues.";
RL Biochemistry 29:1245-1254(1990).
RN [3]
RP SEQUENCE OF 2871-3137 FROM N.A.

```

```

RX MEDLINE=90062147; PubMed=2584214;
RA Bonaldo P., Colombatti A.;
RT "The carboxyl terminus of the chicken alpha 3 chain of collagen VI is
RT a unique mosaic structure with glycoprotein Ib-like, fibronectin type
RT III, and Kunitz modules.";
RL J. Biol. Chem. 264:20235-20239(1989).
CC -I- FUNCTION: Collagen VI acts as a cell-binding protein.
CC -I- SUBUNIT: Trimers composed of three different chains: alpha 1(VI),
CC alpha 2(VI), and alpha 3(VI).
CC -I- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=1;
CC Comment=At least 2 isoforms are produced;
CC Name=1;
CC IsoId=P15989-1; Sequence=Displayed;
CC -I- PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -I- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC -I- SIMILARITY: Contains 1 fibronectin type III domain.
CC -I- SIMILARITY: Contains 12 VWFA domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M24282; AAA03201.1; -.
DR FIR; A37797; A37797.
DR HSPP; P12111; 2KMT.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01391; Collagen; 6.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR Pfam; PF00092; vwa; 11.
DR PRINTS; PR00759; BASICPTASE.
DR PRINTS; PR00453; VWFADOMAIN.
DR ProDom; PD000007; Clg_helix; 1.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00131; KU; 1.
DR SMART; SM00327; VWA; 12.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.
DR PROSITE; PS50234; VWFA; 12.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Cell adhesion; Serine protease inhibitor; Collagen;
KW Signal; Alternative splicing.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 3137 COLLAGEN ALPHA 3(VI) CHAIN.
FT DOMAIN 26 2042 NONHELICAL REGION.
FT DOMAIN 2043 2379 TRIPLE-HELICAL REGION.
FT DOMAIN 2380 3137 NONHELICAL REGION.
FT DOMAIN 38 212 VWFA 1.
FT DOMAIN 241 418 VWFA 2.
FT DOMAIN 444 623 VWFA 3.
FT DOMAIN 644 817 VWFA 4.
FT DOMAIN 842 1014 VWFA 5.
FT DOMAIN 1035 1207 VWFA 6.
FT DOMAIN 1239 1410 VWFA 7.
FT DOMAIN 1441 1621 VWFA 8.
FT DOMAIN 1641 1814 VWFA 9.
FT DOMAIN 1840 2029 VWFA 10.
FT DOMAIN 2407 2587 VWFA 11.
FT DOMAIN 2625 2821 VWFA 12.
FT DOMAIN 2945 3043 FIBRONECTIN TYPE-III.
FT DOMAIN 3068 3137 BPTI/KUNITZ INHIBITOR.
FT SITE 2166 2172 INTERRUPTION IN COLLAGENOUS REGION.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 18:20:26 ; Search time 45 Seconds
(without alignments)
1654.718 Million cell updates/sec

Title: US-09-105-117K-2

Perfect score: 1191
Sequence: 1 VVMEIFITGLIGALLS.....INVVVVVTAIAIKMLMG 236

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rviro.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	29.3	202	16 Q8XKT9	Q8XKT9 Ralstonia s
2	347	29.1	211	16 Q8FE55	Q8FE55 escherichia
3	347	29.1	211	16 Q8Z3W2	Q8Z3W2 salmonella
4	346	29.1	200	16 Q87X28	Q87X28 pseudomonas
5	346	29.1	211	16 Q8ZM68	Q8ZM68 salmonella
6	343	28.8	211	16 Q8XD10	Q8XD10 escherichia
7	343	28.8	211	16 Q7UBP8	Q7UBP8 shigella fl
8	336	28.2	212	16 Q8Y2F1	Q8Y2F1 Ralstonia s
9	335	28.1	205	16 Q8ZHH6	Q8ZHH6 yersinia pe
10	332	27.9	200	16 Q8HW36	Q8HW36 pseudomonas
11	328.5	27.6	204	16 Q9K4K6	Q9K4K6 streptomyces
12	328	27.5	208	16 Q8PHW5	Q8PHW5 xanthomonas
13	326.5	27.4	202	16 Q8UGV8	Q8UGV8 agrobacteri
14	323	27.1	204	16 Q8APD4	Q8APD4 pseudomonas
15	309	25.9	208	16 Q8P6L3	Q8P6L3 xanthomonas
16	302.5	25.4	204	16 Q8YGB7	Q8YGB7 brucella me

17	301	25.3	203	16 Q82P13	Q82P13 streptomyces
18	296	24.9	206	16 Q8ED97	Q8ED97 shewanella
19	291.5	24.5	204	16 Q8G1N8	Q8G1N8 brucella su
20	285	23.9	226	16 Q9CKJ7	Q9CKJ7 pasteurella
21	282.5	23.7	209	16 Q87LL5	Q87LL5 vibrio para
22	281	23.6	186	16 Q83J77	Q83J77 shigella fl
23	275.5	23.1	206	16 Q8DC97	Q8DC97 vibrio vuln
24	266	22.3	211	16 Q7VS78	Q7VS78 bordetella
25	263	22.1	211	16 Q7WQU0	Q7WQU0 bordetella
26	263	22.1	211	16 Q7W1W4	Q7W1W4 bordetella
27	245.5	20.6	211	16 Q9KUN4	Q9KUN4 vibrio chol
28	176.5	14.8	220	16 Q86730	Q86730 bacillus su
29	175.5	14.7	207	16 Q8RHX9	Q8RHX9 fusbacteri
30	173.5	14.6	210	16 Q85420	Q85420 helicobacte
31	170.5	14.3	210	16 Q9ZLC4	Q9ZLC4 helicobacte
32	164	13.8	205	16 Q99V11	Q99V11 staphylococ
33	162.5	13.6	205	16 Q81ET8	Q81ET8 bacillus ce
34	160	13.4	205	16 Q8NX12	Q8NX12 staphylococ
35	158.5	13.3	204	16 Q81RQ2	Q81RQ2 bacillus an
36	158	13.3	193	16 Q88HC2	Q88HC2 pseudomonas
37	157.5	13.2	202	16 Q88X38	Q88X38 lactobacill
38	154.5	13.0	205	16 Q97FY1	Q97FY1 clostridium
39	153.5	12.9	213	16 Q92H04	Q92H04 rickettsia
40	151.5	12.7	208	17 Q87538	Q87538 methanobact
41	148	12.4	200	16 Q9KFP7	Q9KFP7 bacillus ha
42	143	12.0	207	16 Q8LAX1	Q8LAX1 bacillus ce
43	136.5	11.5	197	16 Q8EG90	Q8EG90 shewanella
44	136.5	11.5	206	16 Q8Y2B8	Q8Y2B8 ralstonia s
45	136.5	11.5	213	16 Q913A2	Q913A2 pseudomonas

ALIGNMENTS

RESULT 1

Q8XKT9	PRELIMINARY;	PRT;	202 AA.
ID Q8XKT9	PRELIMINARY;	PRT;	202 AA.
AC Q8XKT9;			
DT 01-MAR-2002 (TREMBLrel. 20, Created)			
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE Probable transmembrane protein.			
GN RSC2024 OR RS03592.			
OS Ralstonia solanacearum (Pseudomonas solanacearum).			
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;			
OC Burkholderiaceae; Ralstonia.			
OX NCBI_TaxID=305;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=GM11000;			
RX MEDLINE=21681879; PubMed=11823852;			
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,			
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,			
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,			
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,			
RA Siguiet P., Thebaud P., Whalen M., Wincker P., Levy M.,			
RA Weissenbach J., Boucher C.A.;			
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";			
RL Nature 415:497-502(2002).			
DR EMBL; AL646067; CAD15726.1; -			
DR GO; GO:0016020; C:membrane; IEA.			
DR GO; GO:0005293; F:lysine permease activity; IEA.			
DR GO; GO:0006865; P:amino acid transport; IEA.			
DR InterPro; IPR001123; LysE.			
DR Pfam; PF01810; LysE; 1.			
KW Complete proteome.			
SQ SEQUENCE 202 AA; 21650 MW; 743460664388E353 CRC64;			

Query Match 29.3%; Score 349; DB 16; Length 202;
Best Local Similarity 35.1%; Pred. No. 1.6e-22;
Matches 80; Conservative 40; Mismatches 72; Indels 36; Gaps 5;

QY 10 GLLIGALLSLLSGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTGLGVLLSNAPI 69

```
Db 7 GILGFSLLAIGSQNAFILRGILGRHVFLGSCVCSALLIVAGVSGFFVLVVKHPY 66
Qy 70 VLDIWRWGIIAYLLWFAVMAAKDAMTKNKEAPQIIEETPTVPDDTPIGGSVAVDTRNR 129
Db 67 IVDVARYGGAFFILWYGI-----SRLLSALRGD 94
Qy 130 VRVEVSVDKQVWV-KEMLMALVTLWLPNAYLDAFVFGVGAQYGDTCGRWTFAGAGA 188
Db 95 SQMDISKAKEADSLPKALLSCIAFTFLNPHVYLDTFVLGSGISAQFGDES-WKFGVGASA 153
Qy 189 ASLIFLPLVFGGAALSRPL-SSPKVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 235
Db 154 ASLIFLPLVFGGAALSRPL-SSPKVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 200

RESULT 2
Q8FE55 PRELIMINARY; PRT; 211 AA.
AC Q8FE55;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Hypothetical protein yggA.
GN YGGA OR C3501.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=2238234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roach P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AB016766; AN81949.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005293; F:lysine permease activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR003016; Lipoyl_BS.
DR InterPro; IPR001123; Lyse.
DR InterPro; IPR004777; Lyse_exporter.
DR Pfam; PF01810; Lyse; 1.
DR TIGRFAMs; TIGR00948; 2a75; 1.
DR PROSITE; PS00189; LIPOYL; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 211 AA; 23173 MW; 5F22D06A3A8FBD73 CRC64;

Query Match 29.1%; Score 347; DB 16; Length 211;
Best Local Similarity 35.7%; Pred. No. 2.6e-22;
Matches 81; Conservative 37; Mismatches 75; Indels 34; Gaps 4;

Qy 7 FITGLIGALLSLTGPNVIVIKQIKREGLIAVLVCLISDVFLFIAGTLGVLISNA 66
Db 5 YFQGLAIGAMILPLGPONAFVWNGQIRQVHIMIALCALSDLVLCAGIFGGSALLMQ 64
Qy 67 APIVLDIWRWGIIAYLLWFAVMAAKDAMTKNKEAPQIIEETPTVPDDTPIGGSVAVD 126
Db 65 SPWLLALVTWGGVAFLLWYGFAGFKTAMSSNIEL----- 98
Qy 127 RNRVRVSVSDKQVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 185
Db 99 -----ASAEVLKQGRW-KIATMLAVTLNPHVYLDTFVLGSGQLDVPFKRW-FALG 151
Qy 186 AFAASLWFLVFGGAALSRPLSSPKVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 232
Db 152 TISASFLWFGGAILLAALWAPRLRTAKQRIINLVVGCWVWVWVWVWVWVWVWVWVWV 198
```

```
RESULT 3
Q823W2 PRELIMINARY; PRT; 211 AA.
AC Q823W2;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Possible membrane transport protein.
GN STY3222 OR T2984.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TY2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyourami V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337 (2003).
DR EMBL; AL627277; CAD02896.1; -.
DR GO; GO:0016844; AAO70536.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005293; F:lysine permease activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR001123; Lyse.
DR InterPro; IPR004777; Lyse_exporter.
DR Pfam; PF01810; Lyse; 1.
DR TIGRFAMs; TIGR00948; 2a75; 1.
KW Complete proteome.
SQ SEQUENCE 211 AA; 23172 MW; 13CB427CABE5A3FC CRC64;

Query Match 29.1%; Score 347; DB 16; Length 211;
Best Local Similarity 33.6%; Pred. No. 2.6e-22;
Matches 79; Conservative 42; Mismatches 80; Indels 34; Gaps 4;

Qy 3 IWEIFITGLIGALLSLTGPNVIVIKQIKREGLIAVLVCLISDVFLFIAGTLGVDL 62
Db 1 MISTYFGVALGAAMILPLGPONAFVWNGQIRQVHIMIALCALSDLVLCAGIFGGS 60
Qy 63 LSNAPIVLDIWRWGIIAYLLWFAVMAAKDAMTKNKEAPQIIEETPTVPDDTPIGGS 122
Db 61 LLMQSPWLLALVTWGGVAFLLWYGFAGFKTAMSSNIEL----- 98
Qy 123 ATDTNRVRVSVSDKQVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 181
Db 99 -----ASAEVWKQGRW-KIATMLAVTLNPHVYLDTFVLGSGQLDVPFKRW- 147
Qy 182 FAAGAFAASLWFLVFGGAALSRPLSSPKVWVWVWVWVWVWVWVWVWVWVWVWVWV 236
Db 148 FALGTISASFLWFGGAILLAALWAPRLRTAKQRIINLVVGCWVWVWVWVWVWVWV 202

RESULT 4
Q87X28
```

```

ID Q87X28 PRELIMINARY; PRT; 200 AA.
AC Q87X28;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Transporter, Lyse family.
GN PSPT04362.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Pedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Winn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA White O., Fraser C., Collier A.;
RT "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016871; AA057812.1; -.
DR TIGR; PSPT04362; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005293; F:lysine permease activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR001123; LyseE.
DR Pfam; PF01810; LyseE; 1.
KW Complete proteome.
SQ SEQUENCE 200 AA; 21339 MW; BB24A1AB402C0880 CRC64;

Query Match 29.1%; Score 346; DB 16; Length 200;
Best Local Similarity 33.0%; Pred. No. 2.9e-22;
Matches 75; Conservative 48; Mismatches 70; Indels 34; Gaps 4;

QY 7 FTGILLGASLLSIGPQNVIVIKQIKREGLIAVLLVCLISDVFLFIAGTIGVLLSNA 66
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
5 YFNGLLIAAGLIAIGTQAFVLAQGLRREHHVPMVLCWCDAIILVAAGVFGLANVLAH 64

QY 67 APVLDIMRWGGIAYLLIFAVMAAKDMTKVAPQIIETEPTVPDTPLGSAVADT 126
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 NPTLLAVRWGGVIFLTUFGQLRRACSR-----OSLEH-----SAAAGT 105

QY 127 RNRVREVSVDKQVWVVKPMLMAIVLTWLPNAYLDAFVFIGVGAGYQGTGRWIFAGA 186
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
106 RSRRTV-----LLSALVATLLNPHVLTDTLLIGSLGAQOSVPGAYV-AGA 150

QY 187 FAASLIWFLVFGAALSRPLSSPKVWMINVAVVMTALAIKMLM 233
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
151 ASASLIWFSALAIAGNALPWLAPATWRMLDVIAVMMVFSVFQLI 197

RESULT 5
Q8ZM68 PRELIMINARY; PRT; 211 AA.
AC Q8ZM68;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Putative Lyse family, amino acid transport protein.
GN YGGA OR STW3066.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

```

```

RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
DR EMBL; AE008841; AAL21941.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005293; F:lysine permease activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR InterPro; IPR001123; LyseE.
DR InterPro; IPR004777; Lys_exporter.
DR Pfam; PF01810; LyseE; 1.
DR TIGRFAMs; TIGR00948; 2a75; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 211 AA; 23220 MW; 64D6FC8F2F1D0F CRC64;

Query Match 29.1%; Score 346; DB 16; Length 211;
Best Local Similarity 33.6%; Pred. No. 3.1e-22;
Matches 79; Conservative 41; Mismatches 81; Indels 34; Gaps 4;

QY 3 IMEIFITGLLIGASLLSIGPQNVIVIKQIKREGLIAVLLVCLISDVFLFIAGTIGVLD 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MISYFQGFALGAAMILPLGPQAFVWQGIIRQYHMLTALLCALSDVLISAGIFGSA 60

QY 63 LSNAAPVLDIMRWGGIAYLLIFAVMAAKDMTKVAPQIIETEPTVPDTPLGSAV 122
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 LLMQSPMLLALVTWGGVAFLLWYFGALKATAMSNLEL----- 98

QY 123 ATDTNRVREVSVDKQVWVVKPMLMAIVLTWLPNAYLDAFVFIGVGAGYQGTGRWI 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
99 -----ASAQVWQGRW-KIATMLAVTLPNPHVLTDTFVVLGSLGGQLAMEPRW- 147

QY 182 FAAGAFAASLIWFLVFGAALSRPLSSPKVWMINVAVVMTALAIKMLMG 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 FALGTISASFLWFGALLAALWAPLRTAKAQRIINILVGVWMLIAFQLAREG 202

RESULT 6
Q8XD10 PRELIMINARY; PRT; 211 AA.
AC Q8XD10;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Orf, hypothetical protein.
GN YGGA OR 24260 OR EC83794.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Ferna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimantata E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
DR EMBL; AE005522; AAG58049.1; -.

```

Dd		65	SPWLLALVTGGGVVFLWYGFAGFTAKSSNIEL-----	98
Qy		127	RNRVRVEVSVDKQRVVVKPMMLMAVLTVLNNAYLDAFVFIGGCAQYG-DTGRWIFPAAG	185
Dd		99	-----ASAERLVKGHW-KITATMLAVTNLPHVYDFTFVLGSLGGQLDVPEKRW-FAUG	151
Qy		186	AFAASLTWFPLVGFGAALSRLPSLPKWRMINNVAVVNTALAIKL	232
Dd		152	TISASFVFFGLAILAWAPRLTAKSRIINLVGCWVFIALQL	198
<hr/>				
RESULT 8				
ID	Q8Y2FI	PRELIMINARY;	PRT; 212 AA.	
AC	Q8Y2FI			
DT	01-MAR-2002 (TrEMBLrel. 20, Created)			
DT	01-JUN-2002 (TrEMBLrel. 20, Last sequence update)			
DT	01-JAN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Probable transmembrane protein.			
GN	RSC0385 OR RS03353.			
OS	Ralstonia solanacearum (Pseudomonas solanacearum).			
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;			
OC	Burkholderiaceae; Ralstonia.			
OX	NCB TaxID=305;			
ON	[1]			
NCBI	SEQUENCE FROM N.A.			
RP	STRAIN=GMI1000;			
RC	MEDLINE=21681879; PubMed=11823852;			
RX	Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,			
RA	Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,			
RA	Chandler M., Choisme N., Claudel-Renard C., Cumac S., Demange N.,			
RA	Gaspin C., Lavie M., Molsan A., Robert C., Saurin W., Schlex T.,			
RA	Siguier P., Thebaut P., Whalen M., Wincker P., Levy M.,			
RA	Weissenbach J., Boucher C.A.;			
RT	"Genome sequence of the plant pathogen Ralstonia solanacearum.";			
RL	Nature 415:497-502(2002).			
DR	EMBL; AL646059; CADL3913.1; --			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0005293; F:lysine permease activity; IEA.			
DR	GO; GO:0006865; P:amino acid transport; IEA.			
DR	InterPro; IPR001123; LyseE.			
DR	Pfam; PF01810; LyseE, 1.			
KW	Complete proteome.			
SQ	SEQUENCE 212 AA; 21896 MW; E29D663C9497EDBB CRG64;			
<hr/>				
Query Match 28.2%; Score 336; DB 16; Length 212;				
Best Local Similarity 31.0%; Pred. No. 2.3e-21;				
Matches 71; Conservative 44; Mismatches 82; Indels 32; Gaps 1;				
Qy	8	ITGLLGASLLLSIGPONVLIVIKQIRKGLTAVILLCLISDVFLFIAGTLGYDLLSNAA	67	
Dd	16	LSGFGLSGLSVIAIGNAYVILRGREYVIGVWLICALCDMALIAGVAGMTLISH	75	
Qy	68	PIVLIDTWKGGIAYLLWFAVAAKDAMTKVKAPOLIETEPTVPDDTPLGGSVAVTDTR	127	
Dd	76	PAWLTAVMAGAFILAYCARAPAAW-----R	103	
Qy	128	NRRVRVEVSVDKQRVVVKPMMLMAVLTVLNNAYLDADFVFIGGCAQYGDTGRWIFPAAGAF	187	
Dd	104	GAERLAQRNGDKASHAQVLASALASILNPVHYDLTVVLVGLAIGGRYPAMPANVAFAGGW	163	
Qy	188	AASLTWFPLVGFGAALSRLPSLPKWRMINNVAVVNTALAIKLMWG	236	
Dd	164	CASILFSLGFGARLEFPVFPANRVLDALICANNWIALTLIMG	212	
<hr/>				
RESULT 9				
ID	Q8ZHH6	PRELIMINARY;	PRT; 205 AA.	
AC	Q8ZHH6;			
DT	01-MAR-2002 (TrEMBLrel. 20, Created)			
DT	01-WAR-2002 (TrEMBLrel. 20, Last sequence update)			

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Putative LysE type translocator (Hypothetical protein).
 GN YP00918 OR Y3305.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Partridge M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holtroyd S., Jagels K., Kariyasek A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM."
 RL J. Bacteriol. 184:4601-4611(2002).
 DR EMBL: AJ414145; CAC89762.1; -;
 DR EMBL: AE013931; AAM6855.1; -;
 DR PIR: AG0112; AG0112.
 DR GO: GO:0016021; C: integral to membrane; IEA.
 DR GO: GO:0005293; F: lysine permease activity; IEA.
 DR GO: GO:0006865; P: amino acid transport; IEA.
 DR InterPro: IPR001123; LysE.
 DR InterPro: IPR004777; Lys_exporter.
 DR Pfam: PF01810; LysE; 1.
 DR TIGRFAMs: TIGR00948; 2a75; 1.
 DR Hypothetical protein; Complete proteome.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 205 AA; 22164 MW; 2965524391474CDA CRC64;

 Query Match 28.1%; Score 335; DB 16; Length 205;
 Best Local Similarity 33.1%; Pred. No. 2.7e-21;
 Matches 80; Conservative 44; Mismatches 70; Indels 48; Gaps 6;

 QY 3 IMEITGLGASLLSTGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTLGVDL 62
 DB 1 MLAVYIHGFILSAAMILPQNVFVNMQGIKQKHLSASICALSDIILICAGIFGSA 60

 QY 63 LSNAPVILDMRWGGIAYLLWF---AVMAKQAMTKVEAPQIIETEPTVPDPLGG 119
 DB 61 LLSRSPLLALVTWGGVAFIMVYWGALMAA-----LAVTWLNPVYLDTFVVLGSLGGQL 141

 QY 120 SAVATD----TNRVRVSVKQKRWVVKPMIAVLVTLNPNAYLDAPFVIGGGAQ-Y 174
 DB 95 DGVASATSVTQGRWRLVTL-----LAVTWLNPVYLDTFVVLGSLGGQL 141

 QY 175 GDTGRWIFAGAFASLIWFLPLVGGAALSRPLSSPKVWRWINVVAVMTALAKML 234
 DB 142 PDIRPW-FALGAVTASTVWFALALAAWLSFMINRPVQRIINIFVGGVMGFIAFOLAR 200

 QY 235 MG 236
 DB 201 QG 202

 RESULT 10
 Q9HW36
 ID Q9HW36 PRELIMINARY; PRT; 200 AA.
 AC Q9HW36;

DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Probable transporter.
 GN PA4365.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL: AE004852; AAG07753.1; -;
 DR PIR: D83100; D83100.
 DR GO: GO:0016020; C: membrane; IEA.
 DR GO: GO:0005489; F: electron transporter activity; IEA.
 DR GO: GO:0005293; F: lysine permease activity; IEA.
 DR GO: GO:0006865; P: amino acid transport; IEA.
 DR GO: GO:0006118; P: electron transport; IEA.
 DR InterPro: IPR002109; Glutaredoxin.
 DR InterPro: IPR001123; LysE.
 DR Pfam: PF01810; LysE; 1.
 DR PROSITE: PS00195; GLUTAREDOXIN; 1.
 KW Complete proteome.
 SQ SEQUENCE 200 AA; 21144 MW; 7C10CB8D8FE36A4 CRC64;

 Query Match 27.9%; Score 332; DB 16; Length 200;
 Best Local Similarity 33.5%; Pred. No. 4.8e-21;
 Matches 77; Conservative 42; Mismatches 77; Indels 34; Gaps 4;

 QY 7 FITGLLGASLLSTGPQNVIVIKQIKREGIIVLVCLISDVFLFIAGTLGVDLSNA 66
 DB 5 YLNGIIVAAGLIIGAQNAFVLAQSLRREHLSVAALCVFCDAVLVSLGVFLAKLLE 64

 QY 67 APIVLDMRWGGIAYLLWFVAVMAKQAMTKVEAPQIIETEPTVPDPLGGSAVATDT 126
 DB 65 NPTLLAARWGGIAFLTWYGLKALLRLR-----PDAL---GNAETGP 105

 QY 127 RNRVRVSVKQKRWVVKPMIAVLVTLNPNAYLDAPFVIGGGAQYDGTGRWIFAGA 186
 DB 106 RSRKAV-----LLAALAVTLNPNVYLDTFVVLGSLGQAQAFG--AYALGA 150

 QY 187 FAASLIWFLPLVGGAALSRPLSSPKVWRWINVVAVMTALAKMLMG 236
 DB 151 ASASLWFFALGALGAWLAPATWRLDLMVAMMLGWAQQLFRG 200

 RESULT 11
 Q9K4K6 PRELIMINARY; PRT; 204 AA.
 AC Q9K4K6
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Putative membrane transport protein.
 GN SC07308 OR SC5F8.18.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);

OC	Xanthomonadaceae; Xanthomonas.
RN	NCBI_TaxID=92829;
TX	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=306 / ATCC 13902 / XV 101;
RX	MEDLINE=20202145; PubMed=12024217;
RA	da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
RA	Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA	Alves I.M.C.G., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA	Camarote G., Cannavan P., Cardozo J., Chambergo F., Ciapina L.P.,
RA	Ciccarelli R.M.B., Coutinho L.I., Cursino-Santos J.R., El-Dorri H.,
RA	Faria J.B.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA	Fomighieri E.P., Franco M.C., Greggio C.C., Gruber A.,
RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA	Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA	Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA	Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.Z.,
RA	Spinola L.A.F., Takita M.A., Tamara R.E., Teixeira E.C., Tezza R.I.D.,
RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA	Serubaj J.C., Kitajima J.P.;
RT	"Comparison of the genomes of two Xanthomonas pathogens with differing
RT	host specificities";
RL	Nature 417:459-463(2002).
DR	EMBL; AB011957; AA037978.1; -
DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0005293; F:lysine permease activity; IEA.
DR	GO; GO:0006865; P:amino acid transport; IEA.
DR	InterPro; IPRO01123; LytB.
DR	Pfam; PF01810; lytB; 1.
KW	Complete proteome.
SQ	SEQUENCE 208 AA; 23223 MW; 61471259ICF2BB7 CRC64;
	Query Match 27.5%; Score 328; DB 16; Length 208;
	Best Local Similarity 29.6%; Pred. No. 1.1e-20;
	Matches 69; Conservative 47; Mismatches 85; Indels 32; Gaps 2;
QY	2 VMIEFTITGLLGASLLISIGPQLVIVKQGKEGILAVLVCLISDVFLTAGTLGD 61 ::: : : : : : : : : : : : : : : : : : : : :
Dd	1 MTVOATVGAFAGAGLTVAIQAQNFAVFTRQGLRQHVLVTTCALGDIALILLGVAGIG 60 : : : : : : : : : : : : : : : : : : : : : : :
QY	62 LLSSNAAPTVIDIMWGITYALLFWFAVMKAADMTNKVEAPQIETEPTVPDDTPLGGSA 121 : : : : : : : : : : : : : : : : : : : : : : :
Dd	61 VLIAQQWSGLQLVRAGAFLAYVGGARRAWR-----GDGA 98 : : : : : : : : : : : : : : : : : : : : : : :
QY	122 VATDTRNRVRVSVVDQRVVWKPMMLVAIVLPINPNAYLDAPFVFGVGAGQGTGRVI 181 ::: : : : : : : : : : : : : : : : : : : : : :
Dd	99 MSPAEQGEDR-----WQRAMLTCIAFTFLNPVHYDTMLLGLSIATHYPGARWA 148 : : : : : : : : : : : : : : : : : : : : : : :
QY	182 FRAGAFASLTWPVLVGRGAALSPLSSPKVKRWVNNAVUTALAIKML 234 : : : : : : : : : : : : : : : : : : : : : : :
Dd	149 FAHGACVASVSFGUGYGARKLQPFVRPRARWLIOGLIANFMPLLVCVLLL 201 : : : : : : : : : : : : : : : : : : : : : : :
RESULT 13	
Q8UGV8	PRELIMINARY; PRT; 202 AA.
ID	Q8UGV8
DC	Q8UGV8
DT	01-JUN-2002 (TREMBLrel. 21, Created)
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE	LytB family transporter.
GN	LYSE OR ATT0927 OR AGC C.1690.
OS	Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC	Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
RN	NCBI_TaxID=176299;
TX	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21608550; PubMed=11743193;
RA	Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA	Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA	Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,

RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kufayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*
 RT C58.",
 RL Science 294:2317-2323 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
 RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT *Agrobacterium tumefaciens* C58.",
 RL Science 294:2323-2328 (2001).
 DR EMBL; AB009057; AL41941.1; -;
 DR PIR; AG2690; AG2690.
 DR PIR; C97472; C97472.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005293; F:lysine permease activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR InterPro; IPR001123; LysE.
 DR Pfam; PF01810; LysE; I.
 KW Complete proteome.
 SQ SEQUENCE 202 AA; 21943 MW; F43E0FF2A834D16 CRC64;
 Query Match 27.4%; Score 326.5; DB 16; Length 202;
 Best Local Similarity 31.6%; Pred. No. 1.5e-20;
 Matches 73; Conservative 42; Mismatches 83; Indels 33; Gaps 4;
 QY 4 MEITFGLGASLLLSIGPQNVLIKQIKREGILAVLVCLISDFLFIAGTGLVDLL 63
 DB 3 IQIFFTGLTGLSLVIAIGQNAFLVKGGLARSHVAVCATCISDALLIMGVGFQRI 62
 QY 64 SNAAPITVLDIMRWGGIAYLLWFVAVMAKDMTNKVEAPOIIEETPTVDDTPLGSSAVA 123
 DB 63 SAIMPALDIPIMRYAGAFLIYWGAKSLYALRSS-EVLSVAERREAS----- 108
 QY 124 TDRNRVRVSVSKQRVWVKPMLMAVLTWLNPNAYLDAFVFGVCAQYGTGTGWI 183
 DB 109 -----LW-OTLAICLALTFLNPHVYLDVTLTGTTISTQPGFEK-TFA 149
 QY 184 AGAFAASLWPLVFGGAALSRPLSSPKVWRVNVVAVVMTALAIKML 234
 DB 150 AGAATOSLLFFESLGYGARWLRDIPEKPSANRILEGVIAITWTAIFKLVW 200
 RESULT 14
 Q88PD4 PRELIMINARY; PRT; 204 AA.
 AC Q88PD4; (T-EMBLrel. 24, Created)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Transporter, LysE family.
 GN PP0916.
 OS *Pseudomonas putida* (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22430300; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,

RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hohnel J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoft A., Tummeler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile *Pseudomonas putida* KT2440.",
 RL Environ. Microbiol. 4:799-808 (2002).
 DR EMBL; AB016777; AAN66541.1; -;
 DR TIGR; PP0916; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005293; F:lysine permease activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR InterPro; IPR001123; LysE.
 DR Pfam; PF01810; LysE; I.
 KW Complete proteome.
 SQ SEQUENCE 204 AA; 21665 MW; 724CFB736E2B12F CRC64;
 Query Match 27.1%; Score 323; DB 16; Length 204;
 Best Local Similarity 32.2%; Pred. No. 3e-20;
 Matches 73; Conservative 47; Mismatches 73; Indels 34; Gaps 5;
 QY 7 FITGLLGLGASLLLSIGPQNVLIKQIKREGILAVLVCLISDFLFIAGTGLVDLLSNA 66
 DB 10 YLNGMLVAFGLMAIGAQNAFLVLAQSLRREHLPVAAICVCDAILVAAGVGLATVLAH 69
 QY 67 APIVLDMRWGGIAYLLWFVAVMAKDMTNKVEAPOIIEETPTVDDTPLGSSAVATDT 126
 DB 70 NPTLLAIARWGGVAVLIYWGAKSLYALRSS-QLSHQQ-----GGGV----- 110
 QY 127 RNRVRVSVSKQRVWVKPMLMAVLTWLNPNAYLDAFVFGVCAQYGTGTGWI 186
 DB 111 RSRV-----LLSALVTLNPHVYLDVTLTGSLGQQSAPGAVV-AGA 155
 QY 187 PAASLWPLVFGGAALSRPLSSPKVWRVNVVAVVMTALAIKML 233
 DB 156 ASASLLWFSTLGAAMLAFLAPLAPATWRLDMLDMVAVMFAVAAQLI 202
 RESULT 15
 Q8P6L3 PRELIMINARY; PRT; 208 AA.
 AC Q8P6L3; (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Membrane transport protein.
 GN YGGA OR XCC2954.
 OS *Xanthomonas campestris* (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A., P.,
 RA Cicarotte G., Camavari F., Cardoso J., Chamargo F., Chapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:15:20 ; Search time 9505 Seconds
(without alignments)
10825.501 Million cell updates/sec

Title: US-09-105-117K-3

Perfect score: 2374

Sequence: 1 agatactcttgaagaaa.....gtacacatttcagcaaatgg 2374

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sw.*
13: gb_un.*
14: gb_vn.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_ma.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_ther.*
33: em_hgt_mus.*
34: em_hgt_pln.*
35: em_hgt_rod.*
36: em_hgt_mam.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgt_hum.*
40: em_hgt_mus.*
41: em_hgt_ther.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	2374	100.0	2374	1	CGLYSEG	X96471 C-Glutamic
C 2	2374	100.0	2374	6	A93933	A93933 Sequence 2
C 3	2374	100.0	333150	1	AP005277	AP005277 Coryneb
C 4	2374	100.0	349980	6	AX127147	AX127147 Sequence
C 5	993	41.8	993	6	AX063767	AX063767 Sequence
C 6	993	41.8	993	6	AX244055	AX244055 Sequence
C 7	870	36.6	870	6	AX123540	AX123540 Sequence
C 8	870	36.6	870	6	BD165657	BD165657 Novel pol
C 9	822	34.6	822	6	AX063771	AX063771 Sequence
C 10	822	34.6	822	6	AX244059	AX244059 Sequence
C 11	818.6	34.5	308650	1	AP005218	AP005218 Coryneb
C 12	798.8	33.6	1771	1	AB083133	AB083133 Coryneb
C 13	724.4	30.5	349535	1	AX248357	AX248357 Coryneb
C 14	711	29.9	711	6	AX643028	AX643028 Sequence
C 15	708	29.8	708	6	AX123539	AX123539 Sequence
C 16	708	29.8	708	6	BD165656	BD165656 Novel pol
C 17	698.6	29.4	1568	6	E54483	E54483 Heat-resist
C 18	696.8	29.4	1083	6	AX813971	AX813971 Sequence
C 19	649	27.3	1083	6	AX123538	AX123538 Sequence
C 20	627	26.4	627	6	BD165655	BD165655 Novel pol
C 21	627	26.4	627	6	AX067087	AX067087 Sequence
C 22	485.2	20.4	993	6	AP003001	AP003001 Mesorhizo
C 23	170	7.2	345783	1	AP003001	AP003001 Mesorhizo
C 24	164.4	6.9	8472	1	AE008076	AE008076 Agrobacte
C 25	164.4	6.9	13051	1	AE009111	AE009111 Agrobacte
C 26	159.4	6.7	302007	1	SC0939132	SC0939132 Streptomy
C 27	140.6	5.9	303550	1	SC0939131	SC0939131 Streptomy
C 28	135.6	5.7	306250	1	SMES591788	SMES591788 Bordetell
C 29	123.8	5.2	347356	1	BX640437	BX640437 Bordetell
C 30	123.8	5.2	348171	1	BX640412	BX640412 Bordetell
C 31	122.8	5.2	14844	1	MBU34849	MBU34849 Mycobacteri
C 32	122.8	5.2	15239	6	I86263	I86263 Sequence 17
C 33	122.8	5.2	306050	1	BX248341	BX248341 Mycobacte
C 34	122.2	5.1	348251	1	BX640423	BX640423 Bordetell
C 35	121.2	5.1	14869	1	AE007056	AE007056 Mycobacte
C 36	121.2	5.1	346186	1	BX842578	BX842578 Mycobacte
C 37	120.8	5.1	298300	1	AP005025	AP005025 Streptomy
C 38	116.4	4.9	11120	1	RFA311775	RFA311775 Rhodococc
C 39	115.2	4.9	298300	1	AP005025	AP005025 Streptomy
C 40	113	4.8	303550	1	SC0939131	SC0939131 Streptomy
C 41	112.6	4.7	10871	1	AE011965	AE011965 Xanthomon
C 42	112.4	4.7	10336	1	AE005530	AE005530 Escherich
C 43	112.2	4.7	12070	1	AB011413	AB011413 Streptomy
C 44	111.2	4.7	10278	1	AE015317	AE015317 Shigella
C 45	111.2	4.7	266658	1	AP002563	AP002563 Escherich

ALIGNMENTS

RESULT 1
CGLYSEG/c
LOCUS C-Glutamicum lysE and lysG genes.
DEFINITION X96471
ACCESSION X96471.1 GI:1729753
VERSION X96471.1
KEYWORDS lysE gene; lysG gene; lysine export regulator protein; lysine exporter protein; lysine governor.
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
AUTHORS Vrljic, M., Salm, H. and Eggeling, L.
TITLE A new type of transporter with a new type of cellular function:

Pred. No. is the number of results predicted by chance to have a

Db 1234 AATATCGAGCAGCATCGCGCGCATTTGGACAAAAGATCAACGCCCAAGTCCCGCGAT 1175

Qy 1201 GAACAAAGAGCGTCAAGAAATTAACACACGACGACGACCGCAATGAGTCTTTCGCCGTT 1260

Db 1174 GAACAAAGAGCGTCAAGAAATTAACACACGACGACGACCGCAATGAGTCTTTCGCCGTT 1115

Qy 1261 AATTCCTTGTGTTAATCACCAGTACATCTTCGGGTCCGATGACAGTAAAGACTGGCCCC 1320

Db 1114 AATTCCTTGTGTTAATCACCAGTACATCTTCGGGTCCGATGACAGTAAAGACTGGCCCC 1055

Qy 1321 CAAAGCAGACCTGTAAATGAGATTTCCATGATCAACATCGTGACCTATGGAAGTACTTA 1380

Db 1054 CAAAGCAGACCTGTAAATGAGATTTCCATGATCAACATCGTGACCTATGGAAGTACTTA 995

Qy 1381 AGTAAATGATGTTCTTAACATGTTTAAATATAGTCTTATGACCCCACTCAACTGGA 1440

Db 994 AGTAAATGATGTTCTTAACATGTTTAAATATAGTCTTATGACCCCACTCAACTGGA 935

Qy 1441 CACTTGTCTCAATCATGATGAAGCAGCTTCGAAGGCCCTCTTACGCCCTTTCAT 1500

Db 934 CACTTGTCTCAATCATGATGAAGCAGCTTCGAAGGCCCTCTTACGCCCTTTCAT 875

Qy 1501 TTCCCTTCGGCGGTGAGTCAGCGGTTAAAGCTTCAGATCATCGTGGTTCGAGTGT 1560

Db 874 TTCCCTTCGGCGGTGAGTCAGCGGTTAAAGCTTCAGATCATCGTGGTTCGAGTGT 815

Qy 1561 GGTATCGCGCAACCAACCGGCCAAAGCAACCGAGCGGTGAGTCTTGTGCAAGCAGC 1620

Db 814 GGTATCGCGCAACCAACCGGCCAAAGCAACCGAGCGGTGAGTCTTGTGCAAGCAGC 755

Qy 1621 GCGGAAATGTTCTGCAAGCAGAACTAAAGCGCACTATCTGACGCGCTTGTCTGA 1680

Db 754 GCGGAAATGTTCTGCAAGCAGAACTAAAGCGCACTATCTGACGCGCTTGTCTGA 695

Qy 1681 AATCCGTTAAACCATCGCATCAACGAGATTCGCTATCCACATGTTCTTCCCGTGT 1740

Db 694 AATCCGTTAAACCATCGCATCAACGAGATTCGCTATCCACATGTTCTTCCCGTGT 635

Qy 1741 CAACGAGTGTGTTCTGGGGTGAGCAACGCTTCACGCTTGCGCTTGAAGATGAAGCGCA 1800

Db 634 CAACGAGTGTGTTCTGGGGTGAGCAACGCTTCACGCTTGCGCTTGAAGATGAAGCGCA 575

Qy 1801 CACATTAATCTCTCGCGCGTGAGATGTTTGAAGCGGTAAACCGGTGAAGCTTAATCC 1860

Db 574 CACATTAATCTCTCGCGCGTGAGATGTTTGAAGCGGTAAACCGGTGAAGCTTAATCC 515

Qy 1861 CGTGGCGGATGTGAAGTAGTAACCTTGAACCATCGCCACTTGGCCATTCGAACCCC 1920

Db 514 CGTGGCGGATGTGAAGTAGTAACCTTGAACCATCGCCACTTGGCCATTCGAACCCC 455

Qy 1921 CTCATTGGGGATGCCCTACATGTTGATGGGAACTAGATTGGGCTGCCATGCCGCTTT 1980

Db 454 CTCATTGGGGATGCCCTACATGTTGATGGGAACTAGATTGGGCTGCCATGCCGCTTT 395

Qy 1981 ACGCTTGGTCCAAAGATGTGCTTCAAGACGTGACCTGGAAGCGGCGGTGATGGTCC 2040

Db 394 ACGCTTGGTCCAAAGATGTGCTTCAAGACGTGACCTGGAAGCGGCGGTGATGGTCC 335

Qy 2041 TGTGGGCGCAGCGGTATCATTTGTCGGTCCGCGGAGGTTTGTGGAGCAATTCG 2100

Db 334 TGTGGGCGCAGCGGTATCATTTGTCGGTCCGCGGAGGTTTGTGGAGCAATTCG 275

Qy 2101 CCGAGGCTTGGTGGGAGCTTCTTCCGAAACCCCAAGCTCTCCCATGCTAAAGCAGG 2160

Db 274 CCGAGGCTTGGTGGGAGCTTCTTCCGAAACCCCAAGCTCTCCCATGCTAAAGCAGG 215

Qy 2161 AGAATGATCTCTCTGATGATACCCATTTGACACACGATGATTTGGACAGATGGCG 2220

Db 214 AGAATGATCTCTCTGATGATACCCATTTGACACACGATGATTTGGACAGATGGCG 155

Qy 2221 CTTGGAATCTAGATCTTAGCTAGACTCAGACGCGCTCTGATGACGATTCGAGG 2280

Db 154 CTTGGAATCTAGATCTTAGCTAGACTCAGACGCGCTCTGATGACGATTCGAGG 95

Qy 2281 ATTGGGCTTTAGTTACTTCTGAAAGGTTCAAGGTTTTTCACTTCTTCGCCCGAGGAA 2340

Db 94 ATTGGGCTTTAGTTACTTCTGAAAGGTTCAAGGTTTTTCACTTCTTCGCCCGAGGAA 35

Qy 2341 TTGGGCGCAGGACAGTAAACACCTTCAGCAATGG 2374

Db 34 TTGGGCGCAGGACAGTAAACACCTTCAGCAATGG 1

RESULT 2

A93933/c

LOCUS 2374 bp DNA linear PAT 22-JAN-2000

DEFINITION Sequence 2 from Patent WO9723597.

ACCESSION A93933

VERSION A93933.1 GI:6742037

SOURCE Corynebacterium glutamicum

ORGANISM Corynebacterium glutamicum

REFERENCE 1 (bases 1 to 2374)

AUTHORS Vrljic,M. and Eggeling,L.

TITLE PROCESS FOR THE MICROBIAL PRODUCTION OF AMINO ACIDS BY BOOSTED ACTIVITY OF EXPORT CARRIERS

JOURNAL Patent: WO 9723597-A 2 03-JUL-1997;

FEATURES

source Location/Qualifiers

1. .2374

/organism="Corynebacterium glutamicum"

/mol_type="unassigned DNA"

/db_xref="taxon:1718"

Query Match 100.0%; Score 2374; DB 6; Length 2374;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGATACTCTTTTGAAGAAACCATGATGACATTTGGCTGACATTTGCTGCTCGAAAGGC 60

Db 2374 AGATACTCTTTTGAAGAAACCATGATGACATTTGGCTGACATTTGCTGCTCGAAAGGC 2315

Qy 61 TCTTTACGTGGTATTTCTTCTACGGTCCAGAGCTTCACAGCGAGCGCGCTGAGTTCA 120

Db 2314 TCTTTACGTGGTATTTCTTCTACGGTCCAGAGCTTCACAGCGAGCGCGCTGAGTTCA 2255

Qy 121 GCGGAGAGCGGCTGCGCGCTTCTGATTCATCAGCGAGCTTATTCATCATTCGTTG 180

Db 2254 GCGGAGAGCGGCTGCGCGCTTCTGATTCATCAGCGAGCTTATTCATCATTCGTTG 2195

Qy 181 GGTGAGGAAACCGGCGGATGACGGTGAAGAACTTTGTCAGTCAGCTGCCCAACATGGTCT 240

Db 2194 GGTGAGGAAACCGGCGGATGACGGTGAAGAACTTTGTCAGTCAGCTGCCCAACATGGTCT 2135

Qy 241 TGGCGTCAATGCTTTTCTCAACATTTGCGAGGCGCTGCTTCACGGAACAAATATTCGATGG 300

Db 2134 TGGCGTCAATGCTTTTCTCAACATTTGCGAGGCGCTGCTTCACGGAACAAATATTCGATGG 2075

Qy 301 AATTCACAGAGGTTCCCGCGCAGCGAGGTAAGTCCCTGCTCAGGGGCAATGTTGAAGT 360

Db 2074 AATTCACAGAGGTTCCCGCGCAGCGAGGTAAGTCCCTGCTCAGGGGCAATGTTGAAGT 2015

Qy 361 GAACAATATTGATATGCTCCGCAAGCTCAATGACATCGCCCGAGAAACGCGGGCAGTCACT 420

Db 2014 GAACAATATTGATATGCTCCGCAAGCTCAATGACATCGCCCGAGAAACGCGGGCAGTCACT 1955

Qy 421 TGGCGAGATGCGGCTTGCATGAGGCTGCTGCGCGAGCAGAGAGTACGCGCGGATACCGT 480

Db 1954 TGGCGAGATGCGGCTTGCATGAGGCTGCTGCGCGAGCAGAGAGTACGCGCGGATACCGT 1895

Qy 481 GACCACTGATTCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

Db 1894 GACCACTGATTCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1835

QY 541 CAACAACTTGGAGTTTCTGACGCGAGTTGGAGCGGATCGATGAGATTTCCACGACGC 600
Db 1834 CAACAACTTGGAGTTTCTGACGCGAGTTGGAGCGGATCGATGAGATTTCCACGACGC 1775
QY 601 CGGCATCAACATTTGGGCGAAGCCGATTTCCAAACCCGCGAAACTAACCCATCAA 660
Db 1774 CGGCATCAACATTTGGGCGAAGCCGATTTCCAAACCCGCGAAACTAACCCATCAA 1715
QY 661 CATCAGTTTGTATGGCCATATCGGTTCATCACTAGTCCACGACGAGTGTATCCAGCGCA 720
Db 1714 CATCAGTTTGTATGGCCATATCGGTTCATCACTAGTCCACGACGAGTGTATCCAGCGCA 1655
QY 721 CACCTTGGGCGTGCACGCGGCGTGCACATGCTGCTGCGCGGAAACCCACGCGGAA 780
Db 1654 CACCTTGGGCGTGCACGCGGCGTGCACATGCTGCTGCGCGGAAACCCACGCGGAA 1595
QY 781 CCAGATCAGGCTTGGCGGAACGCGCGGCGGCGGAAATCCACGCTCGGTGTCGCGTA 840
Db 1594 CCAGATCAGGCTTGGCGGAACGCGCGGCGGCGGAAATCCACGCTCGGTGTCGCGTA 1535
QY 841 TTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 1534 TTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1475
QY 901 GGTTCAGCAGATTTGCATCAACATGCGGCTTTACCCAAACCGGCTGCTTATCGAGCTCAC 960
Db 1474 GGTTCAGCAGATTTGCATCAACATGCGGCTTTACCCAAACCGGCTGCTTATCGAGCTCAC 1415
QY 961 CTCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Db 1414 CTCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1355
QY 1021 ATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
Db 1354 ATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1295
QY 1081 GCGGCTCTTTCGCTGCGCATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
Db 1294 GCGGCTCTTTCGCTGCGCATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1235
QY 1141 AATATCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
Db 1234 AATATCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1175
QY 1201 GAACAAAGAGCGTCAGAAATTAACACACGAGAGAGAACGCAATGAGTCTTTCGCGCTT 1260
Db 1174 GAACAAAGAGCGTCAGAAATTAACACACGAGAGAGAACGCAATGAGTCTTTCGCGCTT 1115
QY 1261 AATTCCTTGTAAATCACCGATATTCGCGGTCGATGGAAGGCGGCGGCGGCGGCGGCGGCGG 1320
Db 1114 AATTCCTTGTAAATCACCGATATTCGCGGTCGATGGAAGGCGGCGGCGGCGGCGGCGGCGG 1055
QY 1321 CAAGAAGAGCGTGAATGAGATTTCCATGATCACCATGCGACCTATGGAAGTACTTA 1380
Db 1054 CAAGAAGAGCGTGAATGAGATTTCCATGATCACCATGCGACCTATGGAAGTACTTA 995
QY 1381 AGTAAATGATTTGTTTAACTATGTTTAAATAGTCTTCAACCCCATCACTGGA 1440
Db 994 AGTAAATGATTTGTTTAACTATGTTTAAATAGTCTTCAACCCCATCACTGGA 935
QY 1441 CACTTCTCTCAATCATGATGAGCGAGCTTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
Db 934 CACTTCTCTCAATCATGATGAGCGAGCTTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 875
QY 1501 TTCCCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1560
Db 874 TTCCCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 815
QY 1561 GGTATCGCGCACCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
Db 814 GGTATCGCGCACCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 755

QY 1621 GCGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 754 GCGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 695
QY 1681 AATCCCGTTAAACATCGCCATCAACGAGATTCGATATCAACATGTTTCTCCCGGTT 1740
Db 694 AATCCCGTTAAACATCGCCATCAACGAGATTCGATATCAACATGTTTCTCCCGGTT 635
QY 1741 CAACGAGTAGCTTCTTGGGCTGGAGCAACGCTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 634 CAACGAGTAGCTTCTTGGGCTGGAGCAACGCTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
QY 1801 CACATTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 574 CACATTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
QY 1861 CTTGGCGGATGTAAGTAGTAGAACTTGGAAACCATGCGGCACTTGGCCATTTGCAACCCC 1920
Db 514 CTTGGCGGATGTAAGTAGTAGAACTTGGAAACCATGCGGCACTTGGCCATTTGCAACCCC 455
QY 1921 CTCATTTGCGGATGCTACATGTTGATGCGAACTAGATTTGGCTGCGATGCGGCTT 1980
Db 454 CTCATTTGCGGATGCTACATGTTGATGCGAACTAGATTTGGCTGCGATGCGGCTT 395
QY 1981 ACGCTTCGCTGCCAAAGATGCTTTCAAGACCGTGAACCTGCGGCGGCTGCGATGCT 2040
Db 394 ACGCTTCGCTGCCAAAGATGCTTTCAAGACCGTGAACCTGCGGCGGCTGCGATGCT 335
QY 2041 TGTGGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 334 TGTGGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 275
QY 2101 CCGAGGCGCTTGGTGGGCGCTTCTTCCGAAACCCCAAGCTGCTCCCATGCTTAAAGCAGG 2160
Db 274 CCGAGGCGCTTGGTGGGCGCTTCTTCCGAAACCCCAAGCTGCTCCCATGCTTAAAGCAGG 215
QY 2161 AGAAGTATCTCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 214 AGAAGTATCTCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 155
QY 2221 CTTGGATCTAGATCTCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 2280
Db 154 CTTGGATCTAGATCTCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 95
QY 2281 ATTGGCGCTTGGTGGGCGCTTCTTCCGAAACCCCAAGCTGCTCCCATGCTTAAAGCAGG 2340
Db 94 ATTGGCGCTTGGTGGGCGCTTCTTCCGAAACCCCAAGCTGCTCCCATGCTTAAAGCAGG 35
QY 2341 TTGGGCGGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2374
Db 34 TTGGGCGGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 3
AP005277
LOCUS
DEFINITION
AP005277 Corynebacterium glutamicum ATCC 13032 DNA, linear BCT 08-AUG-2002
4/10.
AP005277 AP005277 BA000036
AP005277.1 GI:21323710
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum ATCC 13032
Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1
AUTHORS
TITLE
JOURNAL
REFERENCE
2 (bases 1 to 333150)
AUTHORS
TITLE
JOURNAL
Nakagawa, S.
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
Unpublished
Nakagawa, S.
Direct Submission
Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.

Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,
 Tokyo 194-8533, Japan (E-mail: enakagawa@axanagen.com,
 Tel: 81-44-829-3031, Fax: 81-44-813-1651)
 This sequence is conducted by collaboration of Kyowa Hakko Kogyo
 Co. Ltd. And Kitasato University.

FEATURES

source

1. 333150
 /organism="Corynebacterium glutamicum ATCC 13032"
 /mol_type="genomic DNA"
 /strain="ATCC 13032"
 /db_xref="taxon:196627"
 /note="ATCC 13032"

gene

complement(63..1292)
 /gene="Cgl0944"
 /complement(63..1292)
 /gene="Cgl0944"
 /codon_start=1

CDS

/transl_table=11
 /product="Hypothetical membrane protein"
 /protein_id="BAB98337.1"
 /db_xref="GI:21323711"

/translation="MTNPTEERNARLIWANGLQNGDQVAAKTVLPWLLQAAGAP
 FLALLVPIRRAGMLPQAAITGVLRQTSKVMVIGNSGQFVSALGIGVAALPLRG
 WALGITVIVLLAALSFRSCMSASKQVQKVKISKRGKLTGRTATVIGVGMGLVAGL
 AIALFGSHSPTRVLAADVIASSFWLFAISVFARIEYAKPATPKNAFSAHPWRKCI
 AAKDDKAFRRVIRVSMMLVTALSTAFIALAASGNSIDSLGFFLISGLASWGG
 RISGIMSDSSKNVMAGGALFGSIVLIIIVLSSAPAPQINTLVLPFLITIAHTA
 IRVARTYVMDAGDQRTYVADANTLMGVLLVIGALSGFIAFGNEALLFLAAI
 GLIGTISARGLNEVSAG"

gene

1382..2917
 /gene="Cgl0945"
 1382..2917
 /gene="Cgl0945"
 /codon_start=1

CDS

/transl_table=11
 /product="Putative multicopper oxidase"
 /protein_id="BAB98338.1"
 /db_xref="GI:21323712"

/translation="MPELNERTFFKAGVLAATVVGQVLVACSSDDVRVGGEPRTL
 PIPADLGTREGSVHPALEAQGTGSEQLPQVTKWGFNGTHLGTITLVKKGDVHV
 DVNNIDMTVHHGKMLPAIADGQHPSTPGQGTWSTVANDATLVHYPHTG
 LTGLVHAGLAWLVEDEATDKLDPREYGVDDIPLVMDHRELEDSGLDELDPL
 GLGDPRTANGITNAHFDAITRVPRVLVNSNRRFYNLAFSDTTFQVLTASGSLD
 ELQDPTTALGEGGERWEIVLEPEGVDFVLESVDYEDVDEDFVDFGMSDFQL
 LITIGSDSDAQAALPGVILVKSTEDPDIATERTFIMNTFSINDLQMDQVDDVID
 HDQPEVMIYVNDNSDPHNVHDARFKVLKFEQTDVLEFNDGMDQTVGLPGATATL
 AVFEHYPDPQWPMYTHYCHMLYHEDQGMGQFVIVERGDEPAVILGSGSSIDSAGG
 HAH"

gene

3063..3983
 /gene="Cgl0946"
 3063..3983
 /gene="Cgl0946"
 /notes="PF00005:ABC transporter"

CDS

TIGR00960:ja0501902: Type II (General) Secretory Pathway
 (ITSP) Family protein
 TIGR00968:ja0106801: sulfate transport system permease
 protein
 TIGR01166:cbiO: cobalt transport protein ATP-binding
 subunit
 TIGR01184:ntrCD: nitrate transport ATP-binding subunits C
 and D

TIGR01186:prov: glycine betaine/L-proline transport ATP
 binding subunit
 TIGR01187:pota: spermidine/putrescine ABC transporter
 ATP-binding subunit
 TIGR01188:draA: daunorubicin resistance ABC transporter
 ATP-binding subunit
 TIGR01189:cmcA: heme exporter protein CmcA"

gene

/codon_start=1
 /transl_table=11
 /product="ABC-type transporter, ATPase component"
 /protein_id="BAB98339.1"
 /db_xref="GI:21323713"

gene

4003..4836
 /gene="Cgl0947"
 4003..4836
 /gene="Cgl0947"
 /codon_start=1
 /transl_table=11
 /product="ABC-type transporter, permease components"

CDS

/protein_id="BAB98340.1"
 /db_xref="GI:21323714"
 /translation="MTTSHTARGLOHATPERKTSFFKTSLFAEMLQFRNKTLTLM
 ATVPVIGIPLLLFLNGGAESANSFDYFVMTLLFQVTVLSMATTRDERVLAKR
 LRGEARDIIGALCPGALLLITVVIPLLMVIGAPINLVPVFAVLLGLLL
 CSALMTSGFTRNAEAAQSMFVPLAMGSGSIRPVFGDSIVADILATYTPFAALS
 DLVQIGWAGATTADSVGGVEAANFAGIFQDMLIPLGILAAATAAVNAARNYMRWDSY
 R"

gene

4879..6138
 /gene="Cgl0948"
 4879..6138
 /gene="Cgl0948"
 /EC_number="2.7.3.-"
 /codon_start=1
 /transl_table=11
 /product="Two-component system, sensory transduction
 histidine kinases"

CDS

/protein_id="BAB98341.1"
 /db_xref="GI:21323715"
 /translation="WNKDFWTAGTARWFSRGVSLASPVTAPINSRRLPNLAKYTL
 YTRVSIQAIPVLLSAYFLGIVANAGTLNPSFVLLGFSVILLIVLVLYQPSLNS
 YTPRSVQPFETGLVNLGVVSVVLOIPGLNMSDMTRATALIFTLCVLLGIAYI
 PMNRYVWVLIAMSAVLMTSTTVDYLSALWVPLMAGTVLSVMTVDVWKVEVS
 RELEASURVTEERLPAQELHDTLQHLAAMSVKSELALAKGDDRLNELRELQK
 LTRTSMSEMDVVGRTVNLATEIEGAKSLADAHILHSVITGTSQVSPHARELCAM
 LVREATTNLRSDSDATLTLSTSEVRMNNGVNKDITGLSLGSLARSFAESAGMTL
 IVSREDDQFSVRMLINAPANTPAEKA"

gene

6139..6750
 /gene="Cgl0949"
 6139..6750
 /gene="Cgl0949"
 /notes="PF00072:Response regulator receiver domain
 PF00196:Bacterial regulatory proteins, luxR family"

CDS

/codon_start=1
 /transl_table=11
 /product="Two-component system, response regulators
 consisting of a CheY-like receiver domain and a HTH
 DNA-binding domain"
 /protein_id="BAB98342.1"
 /db_xref="GI:21323716"
 /translation="MISITADDEALIASLATLLSLEPDLVPTAGSGEELIETWA
 DPSNRDVCVLQGGIDGIDTATFLMETPDLAVLIVTSHPROLKRALAAGVLG
 FLKSTSDPETAATRTVHAGRYIDPELAAMTISAGESPLTNREVEVLQGLSLA
 EETAAHLAPGTTNYLSQMTKVGQAKRFEATFRAREGLM"

gene

complement(6781..6987)
 /gene="Cgl0950"
 complement(6781..6987)
 /gene="Cgl0950"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"

CDS

/protein_id="BAB98343.1"
 /db_xref="GI:21323717"
 /translation="MKHYWMSYFAGPLNTPHTKVFVMDIVSVNHNFWSTNNWYN
 MTGLPRETFVPAFVGSITVLVLSF"
 7048..8205
 /gene="Cgl0951"
 7048..8205
 /gene="Cgl0951"
 /codon_start=1

Thu Mar 18 12:31:04 2004

129575 ACAGTTCCGTCCTCCAAAGATGTGTTTCAGACCGTACCTGGACGGCGCTCGATGCTCC 129634
 2041 TGTGGGGCCGACGGCGCTATCCATTGTCCCGTTCGGCGGGAAGGTTTGTGAGGCAATTCG 2100
 129635 TGTGGGGCCGACGGCGCTATCCATTGTCCCGTTCGGCGGGAAGGTTTGTGAGGCAATTCG 129694
 2101 CCGAGGCTTGTGTGGGAGCTTCTTCCGGAACCCAGCTGCTCCCATGCTAAAGCAGG 2160
 129695 CCGAGGCTTGTGTGGGAGCTTCTTCCGGAACCCAGCTGCTCCCATGCTAAAGCAGG 129754
 2161 AGAAGTCACTCTCTCGATGAGATACCCATTGACACCGCATGTTATGCAACCGATGCG 2220
 129755 AGAAGTCACTCTCTCGATGAGATACCCATTGACACCGCATGTTATGCAACCGATGCG 129814
 2221 CCGAATCTAGATCTCTAGCTAGCTACAGACCGCTGTTGATGACGCAATCGAGGG 2280
 129815 CCGAATCTAGATCTCTAGCTAGCTACAGACCGCTGTTGATGACGCAATCGAGGG 129874
 2281 ATTGGGGCTTGTGTGAGTCTCTGAAAGGTTTCACTTCTTCCCGCGCAGGAA 2340
 129875 ATTGGGGCTTGTGTGAGTCTCTGAAAGGTTTCACTTCTTCCCGCGCAGGAA 129934
 2341 TTGGGCGAGGAGGAGTAAACACCTTCAGCAATGG 2374
 129935 TTGGGCGAGGAGGAGTAAACACCTTCAGCAATGG 129968

RESULT 5
 AX063767 993 bp DNA linear PAT 24-JAN-2001
 LOCUS Sequence 49 from Patent WO0100843.
 DEFINITION AX063767
 ACCESSION AX063767
 VERSION AX063767.1 GI:12541479
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacterium.
 1
 Pompeius, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.
 corynebacterium glutamicum genes encoding metabolic pathway
 proteins
 Patent: WO 0100843-A 49 04-JAN-2001;
 BASF AKTENGESSELLSCHAFT (DE)
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1. .993
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101_973
 /note="unnamed protein product; RXA01393"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC25125.1"
 /db_xref="GI:12541480"
 /db_xref="REMBL:CAC25125"
 /translation="MNPIQDITLIIIDEGSFEGLASLISFSAVSORVKALEHVHG
 RLVSRTPAKATENGVEVQAAKMWLLQATKAQSGELAEIPLTTAINADSLSTW
 PVPFVNEVASGATLRLLEDAHTLSILRRGDLVCAVTRNANPAGCEVRLGTR
 HLAITATPSLRDAYWVDKLDWAMPVIRFGPKVDLDRDLDGRVDPGVRVRVSVS
 AEGGSAIRGRIGWLLPETOAAAPMLKAGEVILLDEIPIDTPMYWRWRLSRSLRL
 TDAVDAATEGLRP"
 CDS
 1321 CAAAGCAGACCTGTATGAGATTTCATGATCACCATCGTACCTATGGAAGTACTTA 1380
 1 CAAAGCAGACCTGTATGAGATTTCATGATCACCATCGTACCTATGGAAGTACTTA 60

ORIGIN

Query Match 41.8%; Score 993; DB 6; Length 993;
 Best Local Similarity 100.0%; Pred. No. 5.4e-233; Indels 0; Gaps 0;
 Matches 993; Conservative 0; Mismatches 0;
 1321 CAAAGCAGACCTGTATGAGATTTCATGATCACCATCGTACCTATGGAAGTACTTA 1380
 1 CAAAGCAGACCTGTATGAGATTTCATGATCACCATCGTACCTATGGAAGTACTTA 60

901 GGTGACGAGATTGCCATCAACATGGGCTTTTACCACAAACCCGCTGCTTATCGACGCTAC 960
 128495 GGTGACGAGATTGCCATCAACATGGGCTTTTACCACAAACCCGCTGCTTATCGACGCTAC 128554
 961 CTCACCCGACACCGGTTGGGCTGAGTGCACCGCGCAACCGCCCAAGCGCTGTC 1020
 128555 CTCACCCGACACCGGTTGGGCTGAGTGCACCGCGCAACCGCCCAAGCGCTGTC 128614
 1021 ATCGGCGAGGTTGGTCTGTTCTTCAATGATCTGTGGCGCTTCCACTTGTGTTGCTAT 1080
 128615 ATCGGCGAGGTTGGTCTGTTCTTCAATGATCTGTGGCGCTTCCACTTGTGTTGCTAT 128674
 1081 GCGCTCTTTCGCTGCAATGACGCGAACAACATTAACAGGTAAAGGATGACCCACAGCGAT 1140
 128675 GCGCTCTTTCGCTGCAATGACGCGAACAACATTAACAGGTAAAGGATGACCCACAGCGAT 128734
 1141 AATATCGAGCAGATCGCGCGGNTGACAAAGATCAACGCCCAAGGTGCGCGGAT 1200
 128735 AATATCGAGCAGATCGCGCGGNTGACAAAGATCAACGCCCAAGGTGCGCGGAT 128794
 1201 GAACAAAGAGCTGAGAAATTAACACACGAGAGAACCGCAATGAGTCTTTCGCGCTT 1260
 128795 GAACAAAGAGCTGAGAAATTAACACACGAGAGAACCGCAATGAGTCTTTCGCGCTT 128854
 1261 AATTCCTTGTATCAACAGTATCACTTCGCGTCCGATGACAGTAAAGACTGCGGCC 1320
 128855 AATTCCTTGTATCAACAGTATCACTTCGCGTCCGATGACAGTAAAGACTGCGGCC 128914
 1321 CAAAGCAGAGCTGATCAAGATTTCATGATCACCATCGTACCTATGGAAGTACTTA 1380
 128915 CAAAGCAGAGCTGATCAAGATTTCATGATCACCATCGTACCTATGGAAGTACTTA 128974
 1381 AGTAAATGATGTTTCTTAAACATGTTTAAATAGCTTCAAGACCCCACTCACTGGA 1440
 128975 AGTAAATGATGTTTCTTAAACATGTTTAAATAGCTTCAAGACCCCACTCACTGGA 129034
 1441 CACTTGTCTCAATGATGAGAGGAGCTTCAAGGCGCTTCAAGGCGCTTCAAGGCGCTT 1500
 129035 CACTTGTCTCAATGATGAGAGGAGCTTCAAGGCGCTTCAAGGCGCTTCAAGGCGCTT 129094
 1501 TTCCCTCTGGGCTGAGTCAAGCGCTTAAAGCTCTCGAGCATCACTGCGGTGAGTGT 1560
 129095 TTCCCTCTGGGCTGAGTCAAGCGCTTAAAGCTCTCGAGCATCACTGCGGTGAGTGT 129154
 1561 GGTATCGGCAACCCGCGCAAGCAACCGAGCGGTGAGTCTTGTGCAAGCAGC 1620
 129155 GGTATCGGCAACCCGCGCGCAAGCAACCGAGCGGTGAGTCTTGTGCAAGCAGC 129214
 1621 GCGAAATGTTGTTGCTGCAAGCAGCAAACTAAAGCGCAACTATCTGGACGCTTGTCTGA 1680
 129215 GCGAAATGTTGTTGCTGCAAGCAGCAAACTAAAGCGCAACTATCTGGACGCTTGTCTGA 129274
 1681 AATCCGCTTAAACATCGGCACTCAACGAGATTTCGATTCACATGTTTCTTCCCGTGT 1740
 129275 AATCCGCTTAAACATCGGCACTCAACGAGATTTCGATTCACATGTTTCTTCCCGTGT 129334
 1741 CAAAGAGTACTTCTTGGGCTGAGCAACGCTCACTGCGCTTGGAGATGAGCGCA 1800
 129335 CAAAGAGTACTTCTTGGGCTGAGCAACGCTCACTGCGCTTGGAGATGAGCGCA 129394
 1801 CACATATCTTCTGCGCGGTGGAGATGTTTATGAGCGGTAAACCGCTGAAGCTTAATCC 1860
 129395 CACATATCTTCTGCGCGGTGGAGATGTTTATGAGCGGTAAACCGCTGAAGCTTAATCC 129454
 1861 CCGGCGGATGTGAGTATGAGATTGGAACCATGCGCACTTGGCCATTGGCAACCC 1920
 129455 CCGGCGGATGTGAGTATGAGATTGGAACCATGCGCACTTGGCCATTGGCAACCC 129514
 1921 CTCATTGGGAGATCCCTACATGTTGATGGAACCTAGATTGGGCTCGATGCCGCTTCT 1980
 129515 CTCATTGGGAGATCCCTACATGTTGATGGAACCTAGATTGGGCTCGATGCCGCTTCT 129574
 1981 ACGCTTCGTCCTCCAAAGATGTGTTTCAAGACCGTACCTGAGACGGCGCGCTCGATGGTCC 2040

1381 ACTAAATGATGTTGTTCTTACATGTTTAAATATAGTCTCATGAACCCGATTTCAACTGGA 1440
Db |||||
61 AGTAAATGATGTTGTTCTTAAATGTTTAAATATAGTCTCATGAACCCGATTTCAACTGGA 120
Qy |||||
1441 CACTTTGCTCTCAATCATGATGAAGCAGCTTCAAGAGCGCTCTTACGCTTTCAT 1500
Db |||||
121 CACTTTGCTCTCAATCATGATGAAGCAGCTTCAAGAGCGCTCTTACGCTTTCAT 180
Qy |||||
1501 TTCCCTCCGCGGTGAGTCAAGCGGTTAAAGCTCTCAGCATCACTGGTCCGAGTGT 1560
Db |||||
181 TTCCCTCCGCGGTGAGTCAAGCGGTTAAAGCTCTCAGCATCACTGGTCCGAGTGT 240
Qy |||||
1561 GTATTCGCGCACCAACCGCGCAAGCAACCGAGCGGTGAAGTCTTGTGCAAGCAGC 1620
Db |||||
241 GGTATTCGCGCACCAACCGCGCAAGCAACCGAGCGGTGAAGTCTTGTGCAAGCAGC 300
Qy |||||
1621 GCGGAAATGTTGTTCTGCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTGTCTGA 1680
Db |||||
301 GCGGAAATGTTGTTCTGCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTGTCTGA 360
Qy |||||
1681 AATCCCGTTAAACCATGCCATCAACGAGATTCGCTATCCATGTTTCTCCCGTGT 1740
Db |||||
361 AATCCCGTTAAACCATGCCATCAACGAGATTCGCTATCCATGTTTCTCCCGTGT 420
Qy |||||
1741 CAACGAGTAGCTTCTGGGGTGGAGCAACGCTCAAGCTGCGCTTGAAGTGAAGCGCA 1800
Db |||||
421 CAACGAGTAGCTTCTGGGGTGGAGCAACGCTCAAGCTGCGCTTGAAGTGAAGCGCA 480
Qy |||||
1801 CACATATCTCTGCGCGGTGGAGATGTTTATGAGCGGTAAACCGGTGAAGCTAATCC 1860
Db |||||
481 CACATATCTCTGCGCGGTGGAGATGTTTATGAGCGGTAAACCGGTGAAGCTAATCC 540
Qy |||||
1861 CGTGGCGGATGTAAGTAGTAGAATTGGAACTATGCGGCTGCGGCTTGAAGTGAAGCGCA 1920
Db |||||
541 CGTGGCGGATGTAAGTAGTAGAATTGGAACTATGCGGCTGCGGCTTGAAGTGAAGCGCA 600
Qy |||||
1921 CTCAATGCGGATGCTTACATGTTGATGGAACTAGATGCGGCTGCGGCTTGAAGTGAAGCGCA 1980
Db |||||
601 CTCAATGCGGATGCTTACATGTTGATGGAACTAGATGCGGCTGCGGCTTGAAGTGAAGCGCA 660
Qy |||||
1981 ACCTTTGCGTCCCAAGATGCTTCAAGACCGGTGAACCGGTGAAGCTAATCC 2040
Db |||||
661 ACCTTTGCGTCCCAAGATGCTTCAAGACCGGTGAACCGGTGAAGCTAATCC 720
Qy |||||
2041 TGTGGGCGGAGCGGTATCCATGTTCCGCTGCGGAGGTTTGTGAGCAATTCG 2100
Db |||||
721 TGTGGGCGGAGCGGTATCCATGTTCCGCTGCGGAGGTTTGTGAGCAATTCG 780
Qy |||||
2101 CCGAGCGCTTGGTGGGACTTCTCCGAAACCAAGCTGCTCCATGCTTAAAGCAGG 2160
Db |||||
781 CCGAGCGCTTGGTGGGACTTCTCCGAAACCAAGCTGCTCCATGCTTAAAGCAGG 840
Qy |||||
2161 AGAAGTATCTCTCCATGATGATACCAATGACACCGATGATTTGGCAAGTGGCG 2220
Db |||||
841 AGAAGTATCTCTCCATGATGATACCAATGACACCGATGATTTGGCAAGTGGCG 900
Qy |||||
2221 CTTGGATATCTCTCTAGTACTACACGACCGCTGTTGATGAGCAATTCGAGG 2280
Db |||||
901 CTTGGATATCTCTCTAGTACTACACGACCGCTGTTGATGAGCAATTCGAGG 960
Qy |||||
2281 ATTGGCGCTTACTTCTTCTGAAAGGTTTCAAG 2313
Db |||||
961 ATTGGCGCTTACTTCTTCTGAAAGGTTTCAAG 993

RESULT 6
AX244055
LOCUS AX244055 993 bp DNA linear PAT 29-SEP-2001
DEFINITION Sequence 47 from Patent WO0166573.
ACCESSION AX244055
VERSION AX244055.1 GI:15859119
KEYWORDS
SOURCE Corynebacterium glutamicum

ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
1
REFERENCE Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O., Haberhauer, G.,
AUTHORS Kim, J.W., Lee, H.S. and Hwang, B.J. 2001.
TITLE Corynebacterium glutamicum genes encoding metabolic pathway
JOURNAL proteins
PATENT: WO 0166573-A 47 13-SEP-2001;
BASF AKTIENGESELLSCHAFT (DE)
FEATURES
source
1. 993
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101.. 973
/note="unnamed protein product; RXA01393"
/codon_start=1
/transl_table=11
/protein_id="CAC88506.1"
/db_xref="GI:15859120"
/translation="NNPQLDTLLSIIDEGSPFEGASLALSISSAVSORVKALBEHV
RVLSRTPAKATEAGVLAQAKMVLQAKTQALSGRLAIPITTAIINADLSLW
FPPVNEVSWGATLTLRLDEAHTLSLRGDLGAVTREPANVAGVEVLTWR
HLATATSLDADTMDVGLDWAAMPVLRFGPKDVLQDRDLGVDGPGVRRVSLVPS
AEGFGEAIRREGLGWGLLPETQAAPMLKAGEVILLDEIPIDTPMTWWRLESRLRL
TDVVDAAIEGLRP"

ORIGIN

Query Match 41.88; Score 993; DB 6; Length 993;
Best Local Similarity 100.0%; Pred. No. 5.4e-233; Indels 0; Gaps 0;
Matches 993; Conservative 0; Mismatches 0;
1371 CAAAGCAGACCTGTATGAAGATTTCCATGATCACCATCGTGACCTATGAAGTACTTA 1380
Db 1 CAAAGCAGACCTGTATGAAGATTTCCATGATCACCATCGTGACCTATGAAGTACTTA 60
Qy 1381 AGTAAATGATGTTCTTAAACATGTTTAAATATAGTCTCATGAACCCGATTTCAACTGGA 1440
Db 61 AGTAAATGATGTTCTTAAACATGTTTAAATATAGTCTCATGAACCCGATTTCAACTGGA 120
Qy 1441 CACTTTGCTCTCAATCATGATGAAGCAGCTTCAAGAGCGCTCTTACGCTTTCAT 1500
Db 121 CACTTTGCTCTCAATCATGATGAAGCAGCTTCAAGAGCGCTCTTACGCTTTCAT 180
Qy 1501 TTCCCTCCGCGGTGAGTCAAGCGGTTAAAGCTCTCAGCATCACTGGTCCGAGTGT 1560
Db 181 TTCCCTCCGCGGTGAGTCAAGCGGTTAAAGCTCTCAGCATCACTGGTCCGAGTGT 240
Qy 1561 GGTATTCGCGCACCAACCGCGCAAGCAACCGAGCGGTGAAGTCTTGTGCAAGCAGC 1620
Db 241 GGTATTCGCGCACCAACCGCGCAAGCAACCGAGCGGTGAAGTCTTGTGCAAGCAGC 300
Qy 1621 GCGGAAATGTTGTTCTGCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTGTCTGA 1680
Db 301 GCGGAAATGTTGTTCTGCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTGTCTGA 360
Qy 1681 AATCCCGTTAAACCATGCCATCAACGAGATTCGCTATCCATGTTTCTCCCGTGT 1740
Db 361 AATCCCGTTAAACCATGCCATCAACGAGATTCGCTATCCATGTTTCTCCCGTGT 420
Qy 1741 CAAAGTAGTACTTCTGGGTGAGCAACGCTCAGCTCGCTGCGGTGGAAGTGAAGCGCA 1800
Db 421 CAAAGTAGTACTTCTGGGTGAGCAACGCTCAGCTCGCTGCGGTGGAAGTGAAGCGCA 480
Qy 1801 CACATATCTCTGCGCGGTGGAGATGTTTATGAGCGGTAAACCGGTGAAGCTAATCC 1860
Db 481 CACATATCTCTGCGCGGTGGAGATGTTTATGAGCGGTAAACCGGTGAAGCTAATCC 540
Qy 1861 CGTGGCGGATGTAAGTAGTAGAATTGGAACTATGCGGCTGCGGCTTGAAGTGAAGCGCC 1920
Db 541 CGTGGCGGATGTAAGTAGTAGAATTGGAACTATGCGGCTGCGGCTTGAAGTGAAGCGCC 600

1661 CTATCTGACGCGCTTGTGTAATCCCGTTAAACCATCGCATCAACGAGATTCGCTATCC 1720
 241 CTATCTGACGCGCTTGTGTAATCCCGTTAAACCATCGCATCAACGAGATTCGCTATCC 300
 1721 ACATGTTCTCTCCCGTTCACGAGGTAGCTTCTTGGGGTGGAGCAAGCTCAGCGTG 1780
 301 ACATGTTCTCTCCCGTTCACGAGGTAGCTTCTTGGGGTGGAGCAAGCTCAGCGTG 360
 1781 CGCTTGAAGATGAAGCGCACATTTATCTTCTGCGGGTGGAGATGTTTGAAGGCG 1840
 361 CGCTTGAAGATGAAGCGCACATTTATCTTCTGCGGGTGGAGATGTTTGAAGGCG 420
 1841 GTAAACCGGTGAAGCTTAATCCGTTGGGGTGGAGATGTTTGAAGGCG 1900
 421 GTAAACCGGTGAAGCTTAATCCGTTGGGGTGGAGATGTTTGAAGGCG 480
 1901 CACTTGGCCATTGCAACCCCTCATTTGCGGGATGCTTACATGTTGATGGGAACTAGAT 1960
 481 CACTTGGCCATTGCAACCCCTCATTTGCGGGATGCTTACATGTTGATGGGAACTAGAT 540
 1961 TGGGCTCGCATGCGCGTCTTACGCTTCCGTTCCCAAGATGTTTCAAGCGTGACCTG 2020
 541 TGGGCTCGCATGCGCGTCTTACGCTTCCGTTCCCAAGATGTTTCAAGCGTGACCTG 600
 2021 GACGCGCGCTGATGCTTCTGTTGGGGCGGCGGCTTATCCATGTTCCCGGAACTAGAT 2080
 601 GACGCGCGCTGATGCTTCTGTTGGGGCGGCGGCTTATCCATGTTCCCGGAACTAGAT 660
 2081 GGTGTTGTCGAGCAATTCGCGGAGGCTTGTGTTGGGAGCTTCTCCGAAACCCAGCT 2140
 661 GGTGTTGTCGAGCAATTCGCGGAGGCTTGTGTTGGGAGCTTCTCCGAAACCCAGCT 720
 2141 GCTCCCATGCTTAAAGCAGGAGATGATCTTCTCGATGATGATGATGATGATGATGATGAT 2200
 721 GCTCCCATGCTTAAAGCAGGAGATGATCTTCTCGATGATGATGATGATGATGATGATGAT 780
 2201 ATGTTATGCGCAAGATGCGCGCTGGAATCTAGATCTCTAGCTAGCTAGCTAGCTAGCTAGCT 2260
 781 ATGTTATGCGCAAGATGCGCGCTGGAATCTAGATCTCTAGCTAGCTAGCTAGCTAGCTAGCT 840
 2261 GTTGATGCGCAAGATGCGCGCTGGAATCTAGATCTCTAGCTAGCTAGCTAGCTAGCTAGCT 2320
 841 GTTGATGCGCAAGATGCGCGCTGGAATCTAGATCTCTAGCTAGCTAGCTAGCTAGCTAGCT 900

RESULT 8
 BD165657 870 bp DNA linear PAT 17-JAN-2003
 LOCUS BD165657
 DEFINITION Novel polynucleotide.
 ACCESSION BD165657
 VERSION BD165657.1 GI:27871469
 KEYWORDS JP 2002191370-A/3456
 SOURCE unidentified
 ORGANISM unclassified
 1 (bases 1 to 870)
 REFERENCES
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotide
 JOURNAL Patent: JP 2002191370-A 3456 09-JUL-2002;
 KYOWA HAKKO KOGYO CO LTD
 COMMENT OS Corynebacterium glutamicum
 PN JP 2002191370-A/3456
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI, KEIKO OCHIAI,
 PI HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO OZAKI
 PI PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC C12N1/15,
 PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/

1921 CTCTTGGGCGGATGCTTACATGTTGATGGGAACTAGATGCGGCTCGATGCGCGCTTT 1980
 601 CTCTTGGGCGGATGCTTACATGTTGATGGGAACTAGATGCGGCTCGATGCGCGCTTT 660
 1981 ACCTTGGGCGGATGCTTACATGTTGATGGGAACTAGATGCGGCTCGATGCGCGCTTT 2040
 661 ACCTTGGGCGGATGCTTACATGTTGATGGGAACTAGATGCGGCTCGATGCGCGCTTT 720
 2041 TGTGGGCGGATGCTTACATGTTGATGGGAACTAGATGCGGCTCGATGCGCGCTTT 2100
 721 TGTGGGCGGATGCTTACATGTTGATGGGAACTAGATGCGGCTCGATGCGCGCTTT 780
 2101 CCGAGCGCTTGGTGGGCGCTTCTTCCGAAACCCAGCTGCTCCCATGCTTAAAGCGG 2160
 781 CCGAGCGCTTGGTGGGCGCTTCTTCCGAAACCCAGCTGCTCCCATGCTTAAAGCGG 840
 2161 AGAAGTGATCTCTCGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 841 AGAAGTGATCTCTCGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 2221 CCGAGCGCTTGGTGGGCGCTTCTTCCGAAACCCAGCTGCTCCCATGCTTAAAGCGG 2280
 901 CCGAGCGCTTGGTGGGCGCTTCTTCCGAAACCCAGCTGCTCCCATGCTTAAAGCGG 960
 2281 ATTGGCGCTTGGTGGGCGCTTCTTCCGAAACCCAGCTGCTCCCATGCTTAAAGCGG 2340
 961 ATTGGCGCTTGGTGGGCGCTTCTTCCGAAACCCAGCTGCTCCCATGCTTAAAGCGG 1000

RESULT 7
 AX123540 870 bp DNA linear PAT 11-MAY-2001
 LOCUS AX123540
 DEFINITION Sequence 3456 from Patent EP1108790.
 ACCESSION AX123540
 VERSION AX123540.1 GI:14041028
 KEYWORDS Corynebacterium glutamicum
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 1
 REFERENCES
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotide
 JOURNAL Patent: EP 1108790-A 3456 20-JUN-2001;
 KYOWA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source Location/Qualifiers
 1..870
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"

ORIGIN
 Query Match 36.6%; Score 870; DB 6; Length 870;
 Best Local Similarity 100.0%; Pred. No. 9.5e-203;
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1421 ATGAACCCATCACTGAGACCTTGTCTTCAATATTGATGAAGGAGCTTGAAGGC 1480
 1 ATGAACCCATCACTGAGACCTTGTCTTCAATATTGATGAAGGAGCTTGAAGGC 60
 1481 GCTTCTTATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1540
 61 GCTTCTTATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1600
 1541 CATCAGTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1660
 121 CATCAGTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 1601 GAGTCTTGTGCAAGCAGCGGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1660
 181 GAGTCTTGTGCAAGCAGCGGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240

PC 04, C12P13/08.
 PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
 G01N33/566,
 PC G01N33/569, G01N33/68, G01N37/00//C12P21/09, (C12N1/21, C12R1.15),
 PC (C12N1/21, C12R1.13), (C12N1/21, C12R1.01), (C12P13/08, C12R1.15),
 PC C12N15/00,
 PC C12N5/00, C12N15/00
 CC Novel polynucleotide
 FH Key Location/Qualifiers
 FT source 1..870
 FT 1..870 /organism="Corynebacterium glutamicum".
 FEATURES
 source
 1..870
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 36.6%; Score 870; DB 6; Length 870;
 Best Local Similarity 100.0%; Pred. No. 9.5e-203;
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1421 ATGAACCCCAATCAACTGACACTTTGCTCTCAATCATGATGAAGCAGCTTCGAAGGC 1480
 DB 1 ATGAACCCCAATCAACTGACACTTTGCTCTCAATCATGATGAAGCAGCTTCGAAGGC 60
 QY 1481 GCCTCTAGCCCTTCCATTTCCTCCCTCGGGGTGAGTCAAGCGCTTAAAGCTCTCGAG 1540
 DB 61 GCCTCTAGCCCTTCCATTTCCTCCCTCGGGGTGAGTCAAGCGCTTAAAGCTCTCGAG 120
 QY 1541 CATACAGTGGGTGAGTGTGATTCGCGCACCCACCGCCCAAGCCCAAGCGGGT 1600
 DB 121 CATACAGTGGGTGAGTGTGATTCGCGCACCCACCGCCCAAGCCCAAGCGGGT 180
 QY 1601 GAAGTCTTGTGAACGACGCGGAAATGTTGCTGCAAGCAGAACTAAAGCGCAA 1660
 DB 181 GAAGTCTTGTGAACGACGCGGAAATGTTGCTGCAAGCAGAACTAAAGCGCAA 240
 QY 1661 CTATCTGAGCGCTTCTGTAATCCGTTAAACATGCGCATCAACGAGATTGCTATCC 1720
 DB 241 CTATCTGAGCGCTTCTGTAATCCGTTAAACATGCGCATCAACGAGATTGCTATCC 300
 QY 1721 ACATGCTTCTCCCGTGTCAACGAGTACTTCTTGGGTGAGCAACGCTCAACGCTG 1780
 DB 301 ACATGCTTCTCCCGTGTCAACGAGTACTTCTTGGGTGAGCAACGCTCAACGCTG 360
 QY 1781 CCCTTGAAGATGAAGCGCACACATATCTCTGCTGCGGCGTGGAGATGTTTAGAGCG 1840
 DB 361 CCCTTGAAGATGAAGCGCACACATATCTCTGCTGCGGCGTGGAGATGTTTAGAGCG 420
 QY 1841 GTAACCGTGAAGCTAATCCGTTGGCGGATGTAAGTAGTAGAATTGGAAACATGGCG 1900
 DB 421 GTAACCGTGAAGCTAATCCGTTGGCGGATGTAAGTAGTAGAATTGGAAACATGGCG 480
 QY 1901 CACTTGGCCATTCGAACCCCTCAATTCGGGATGCTACATGTTGATGGGAACTAGAT 1960
 DB 481 CACTTGGCCATTCGAACCCCTCAATTCGGGATGCTACATGTTGATGGGAACTAGAT 540
 QY 1961 TGGGTGCGATGCCCGTCTTACGCTTCGFTCCCAAGATGTGCTTCAAGACCGTGACCTG 2020
 DB 541 TGGGTGCGATGCCCGTCTTACGCTTCGFTCCCAAGATGTGCTTCAAGACCGTGACCTG 600
 QY 2021 GACGGGCGCTGATGGTCTGTTGGCGGACGCGGTATCCATTTGCTCCGTCGCGGAA 2080
 DB 601 GACGGGCGCTGATGGTCTGTTGGCGGACGCGGTATCCATTTGCTCCGTCGCGGAA 660
 QY 2081 GCTTTTGTGAGCAATTCGCGAGGCTTGTGTTGGGACTTCTTCCCAACCCAGCT 2140
 DB 661 GCTTTTGTGAGCAATTCGCGAGGCTTGTGTTGGGACTTCTTCCCAACCCAGCT 720
 QY 2141 GCTCCCATGCTAAAGCAGGAGAGTATCTCTCGATGAGATACCCATTGACACACCG 2200
 DB 721 GCTCCCATGCTAAAGCAGGAGAGTATCTCTCGATGAGATACCCATTGACACACCG 780

QY 2201 ATGTATTGGCAACGATGGCGCTTGAATCTAGATCTCTAGCTAGACTCACAGAGCGGTC 2260
 DB 781 ATGTATTGGCAACGATGGCGCTTGAATCTAGATCTCTAGCTAGACTCACAGAGCGGTC 840
 QY 2261 GTTGTATGCAAGCAATCGAGGATTTGGCGGCT 2290
 DB 841 GTTGTATGCAAGCAATCGAGGATTTGGCGGCT 870
 RESULT 9
 AX063771/c
 LOCUS AX063771 Sequence 53 from Patent WO0100843. DNA linear PAT 24-JAN-2001
 DEFINITION AX063771
 ACCESSION AX063771
 VERSION AX063771.1 GI:12541483
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 1
 Pompeius, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.
 corynebacterium glutamicum genes encoding metabolic pathway
 proteins
 JOURNAL Patent: WO 0100843-A 53 04-JAN-2001;
 BASF AKTIBIOTICSELLSCHAFT (DE)
 FEATURES
 Location/Qualifiers
 source
 1..822
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101..802
 /note="unnamed protein product; RXA01394"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC25127.1"
 /db_xref="GI:12541484"
 /translation="MEIFITGLIGASLLLSIGPONVLVIKQIKREGILAVLLVCLII
 SDVFLFTAGTGLVGLLSNAIPVLDIMRWGGIAYLLFAVMAAKDAMTKVNPAYLIE
 ETEFLTPDDTPLGGSAVATDTRNRVSVKORVVKPMLMAVILVTLNPNAYLDA
 FVFVGGVQYQYDTRWIFRAGAPAAASLLIFPLVGFAGAAALSRLSSPKWRHNVVV
 AVMTALAIKLMNG"
 ORIGIN
 Query Match 34.6%; Score 822; DB 6; Length 822;
 Best Local Similarity 100.0%; Pred. No. 6.1e-191;
 Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 629 GATTCGCAAAACCCGCAAACTAACCCATCAATCAGTTTGTATGGCCATCGGTCATC 688
 DB 822 GATTCGCAAAACCCGCAAACTAACCCATCAATCAGTTTGTATGGCCATCGGTCATC 763
 QY 689 ACACTGCTCCAGCAGCTTGTATCCGCGGCACACCTTGGGCTTGGACGCGGGTGCAC 748
 DB 762 ACACTGCTCCAGCAGCTTGTATCCGCGGCACACCTTGGGCTTGGACGCGGGTGCAC 703
 QY 749 AATGCTGCTGCGCGCAAAACCCACACGAGGGAACACAGATCAGGCTTGGCGCAACGCGCA 808
 DB 702 AATGCTGCTGCGCGCAAAACCCACACGAGGGAACACAGATCAGGCTTGGCGCAACGCGCA 643
 QY 809 GCGGCAAAATTCACCGTCCGCTGTGCGCGTATTTGCGCGCGCAGCGCGCGGATAAACACA 868
 DB 642 GCGGCAAAATTCACCGTCCGCTGTGCGCGTATTTGCGCGCGCAGCGCGCGGATAAACACA 583
 QY 869 AACGCTCCAAATACCATTCGGGTTCACACGCTCAGCAGATTCGATCAACATGGGC 928
 DB 582 AACGCTCCAAATACCATTCGGGTTCACACGCTCAGCAGATTCGATCAACATGGGC 523
 QY 929 TTTTACCCAAACCCGCTGCTTATTCAGCGCTCACTCCACCGCAGCGGTTGGCGGCTGCA 988
 DB 522 TTTTACCCAAACCCGCTGCTTATTCAGCGCTCACTCCACCGCAGCGGTTGGCGGCTGCA 463

989	GTGGCCACCGCCGACCGCCCAAGCGCTGTCATCGGCGACGGTGGTTCTGTCTTCTTCA	1048
990		
991		
992		
993		
994		
995		
996		
997		
998		
999		
1000		
1001		
1002		
1003		
1004		
1005		
1006		
1007		
1008		
1009		
1010		
1011		
1012		
1013		
1014		
1015		
1016		
1017		
1018		
1019		
1020		
1021		
1022		
1023		
1024		
1025		
1026		
1027		
1028		
1029		
1030		
1031		
1032		
1033		
1034		
1035		
1036		
1037		
1038		
1039		
1040		
1041		
1042		
1043		
1044		
1045		
1046		
1047		
1048		
1049		
1050		
1051		
1052		
1053		
1054		
1055		
1056		
1057		
1058		
1059		
1060		
1061		
1062		
1063		
1064		
1065		
1066		
1067		
1068		
1069		
1070		
1071		
1072		
1073		
1074		
1075		
1076		
1077		
1078		
1079		
1080		
1081		
1082		
1083		
1084		
1085		
1086		
1087		
1088		
1089		
1090		
1091		
1092		
1093		
1094		
1095		
1096		
1097		
1098		
1099		
1100		
1101		
1102		
1103		
1104		
1105		
1106		
1107		
1108		
1109		
1110		
1111		
1112		
1113		
1114		
1115		
1116		
1117		
1118		
1119		
1120		
1121		
1122		
1123		
1124		
1125		
1126		
1127		
1128		
1129		
1130		
1131		
1132		
1133		
1134		
1135		
1136		
1137		
1138		
1139		
1140		
1141		
1142		
1143		
1144		
1145</		

RESULT 10	AX244059/c	822 bp	DNA	linear	PAT 28-SEP-2001
LOCUS	Sequence 51 from Patent WO0166573.				
DEFINITION	AX244059				
ACCESSION	AX244059.1	GI:15859123			
VERSION					
KEYWORDS					
SOURCE	Corynebacterium glutamicum				
ORGANISM	Corynebacterium glutamicum				
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
	Corynebacterineae; Corynebacteriaceae; Corynebacterium.				
REFERENCE	1				
AUTHORS	Pompejus, M., Kroege, B., Schroeder, H., Zelder, O., Haberhauer, G., Klim, J. W., Lee, H. S. and Hwang, B. J.				
TITLE	Corynebacterium glutamicum genes encoding metabolic pathway				
JOURNAL	Patent: WO 0166573-A 51 13-SEP-2001;				
FEATURES	Source				
	Location/Qualifiers				
	1..822				
	/organism="Corynebacterium glutamicum"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:1718"				
	101..802				
	/note="unamed protein product; RXA01394"				
	/codon_start=1				
	/transl_table=1				
	/protein_id="CAC88508.1"				
	/db_xref="GI:15859124"				
	/db_xref="REMTREMBL:CAC88508"				
	/translation="MFTFTTGLGLASLLTSGPNVLVIKQIKREGILAVLVLCIL SDVFLIAGTGVQLLSNAPILVLDIMRWGGIYALLFAVMAAKDAMTKVAPQIIE ETFEPTVDDTDLGGSAVATDRENRVRVSDKORVWVEMLMALVLTWLPNAYLDA FVFVGGVAGTQDGRWTFPAAGAFASLIWFPLVGFGLAAALSRPLSPKWRINWIV AVVMTALATKLMIMG"				
CDS					
ORIGIN					
Query Match	34.6%	Score 822;	DS 6;	Length 822;	
Best Local Similarity	100.0%;	Pred. No. 6.1e-191;			

Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.
 Comparative Complete Genome Sequence Analysis of the Amino Acid Replacements Responsible for the Thermostability of *Corynebacterium efficiens*
 Genome Res. 13 (7), 1572-1579 (2003)
 22723752
 12840036
 2 (bases 1 to 308650)
 Kawarabayashi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.
 Direct Submission
 Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bionote.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)
 Kawarabayashi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan
 Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
 Itoh, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan
 Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan
 Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan
 The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.
 Location/Qualifiers
 1..308650
 /organism="Corynebacterium efficiens YS-314"
 /mol_type="genomic DNA"
 /strain="YS-314"
 /db_xref="taxon:196164"
 complement(83)..241)
 /note="CE1178"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC17989.1"
 /db_xref="GI:23493017"
 /translation="MTTGIILGTGFGDAGLKGSDIPRGVGIHFRKPPVALQYR
 TGAGARNQ"
 365..742
 /note="CE1179"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC17989.1"
 /db_xref="GI:23493018"
 /translation="WEKSQLOSKTQDFASVDHLSDDAAAFIDNLSRGAMHRAFLH
 LVHCACRQELARQDIDYDVRTECRNDNVTPASDLKAKLSLANECAPGPGAEDAAI
 QRPESFVAKVSVRAVRRTQGR"
 774..1271
 /note="CE1180, similar to A0292256-1|CAC21724.1| percent
 identity: 26 in 155 aa"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC17990.1"
 /db_xref="GI:23493019"
 /translation="WFSNVGWGRIILLFIVGLIVIGPERIPRIMEDLKAALVAARTAI
 NNAQNLDIDGEEDEIRKPMQAIAIRSMSPKTAITRALFDGDDSLFDSFDPKKIM
 ADDTAGARHQVAKTGTTEHAEVVERPADRIDPTQSPNTFGKNDGKPPAGGYT
 WTDLI"
 complement(1353)..2480)
 /note="CE1181, similar to AL390975-15|CAC01353.1| percent
 identity: 63 in 370 aa"
 /codon_start=1
 /transl_table=11
 /product="putative ATP-binding protein Mrp"
 /protein_id="BAC17991.1"

/db_xref="GI:23493020"
 /translation="MMSQVTEASVRSALSRVEDPEIGKPIITELGMKSVIAIDGSDVQV
 EYVITIAACMKTIIVSNTETAIKIDIEGVDGVQVITVMSDEQRRLVSRGNTADP
 VIPPAQPGSTRVYVAVSGKGVKMSIAHFVGNAPVVRGPGMLHRAIQDILGHSVFGM
 LGSDARHQVDDMPQAHGVKMSIAHFVGNAPVVRGPGMLHRAIQDILGHSVFGM
 DUDILLDLPFGTDIAITIAQLPNAELIIVTPQAAADVAERAGSISLQTNKVA
 GVIENMAMVPGDSTMEVFGGGKIAQRISMLTGEDVPVLSVPLDPSLRVGVDV
 GNPIAISEPSTTAANAIAIDTLAKKKSIVGKTLGLGVK"
 complement(2513)..3271)
 /note="CE1182"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC17992.1"
 /db_xref="GI:23493021"
 /translation="MTTATPWADAFDHVWDLFQILEVDHTERVAHLQYFATDRT
 GMYTGSIFNLGGGDAQVNEITRDDLLSLSDLPVPGVQLVLEPPEDAQSI
 DLEVVYTDYATVELPYDLEQIRGLASIPDVGLASKAAHLDAGNMLWTALRRNG
 FETRVSKLIARRPKLPFIPIFSAIKQLNTNSGVGFNFHTVLRADKALSRLKGI
 RTEADQSGNALADISTIRFEDVVVVMADHKRT"
 complement(3282)..3836)
 /note="CE1183, similar to AE007003-3|AAK45527.1| percent
 identity: 57 in 159 aa"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC17993.1"
 /db_xref="GI:23493022"
 /translation="MADYSRLDTPLVGTRRRRKFDDDTVGAIAEKVAFRFTQGY
 LFMQITFVIMNLNIGGAWNDPPTILNLAFTQAAVAPLILLAQNRQEDDK
 STILADRRARETAKDETELRELAVELAGVEMVTRDYLRHLDLGRFLERLEAKL
 DDAAVRNDPHDLNAPTQKQVD"
 complement(3829)..5124)
 /note="CE1184, similar to AE007003-4|AAK45528.1| percent
 identity: 52 in 417 aa"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC17994.1"
 /db_xref="GI:23493023"
 /translation="MRGITRLVAGRLAGVVRGPDITDVIQRVDRVVVITRSGQDSRV
 LGVAVLNKKRIFLPMLRIAIIPKEVNLISGVSILASFKARTGELQVTDIGTK
 VHTDPELQHLGRAVIADELETRDVIKSVAILGERKFRGRTTYIAPWSH
 VHGTAGGETNQTAELIAGFEDMRADVAKELGLPKPQDVVAGLELMDPEASL
 QELSDRQABLTLDIERAADVLEMPDDADLIGELPEPKAGVLELMDPEASL
 VRLMGFNPDTVGMATPEELTMDPSTTVAELAMAKNPDLTSLASLVFVVRPTAT
 PTGKYLGCYVHLKREPPSLVGGILDPPLPLYANDSQETAAFFATYTNLVCGPVL
 DENGHLIGAVAVDILLDMLPDRDAGIRPQVKEETLG"
 5264..5782
 /note="CE1185, similar to AE007003-6|AAK45530.1| percent
 identity: 60 in 155 aa"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC17995.1"
 /db_xref="GI:23493024"
 /translation="MSMQATGSGNNRRELGIKKRPSGVPVSGFOTYBEAORAVDMUS
 DNEPFAELTIIVGDLMEVETVTRLTWGRVLGAGASGAMLGLEGLVGMGLNLF
 SLLTGLVGVVGYVLAAPVYVWASGRKDFTSATQIVAGRYDILCRDPRAREARDMI
 ARNMGLGHPQ"
 5805..6500
 /note="CE1186"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC17996.1"
 /db_xref="GI:23493025"
 /translation="MRTGVKKKLLGLVAVLGGSLVACGSDVDPNRPFPHPQVABEQK
 ENLVAVDDPDYIGHNVISELYKTVLNGRFAVIRHEDNLIADPDLNGETSLVIG
 CTGALLEVNPVLAGELEEEYLAAVAGETDINGEWRDRTFDALLGISLPELTDAP
 SNAQCGFANGLPQNVIPYINPTFRADREILNWSAGIITTSLELQELIEAKELGRS
 RELVKEFVASKG"
 6660..7373

```

/notes="CE1187, similar to AE007262-3|AAK65408.1| percent
identity: 32 in 235 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17997.1"
/db_xref="GI:23493026"
/translation="METTILGIVLLFVAACVQGAIFGFGMLAAPIALMRPDLPLGLI
LLAFGLSLATWARDREVENPVWMSLGRIPGSLVGTWVAVMPVAGLKIVLATV
ILGLTSLVWSPGHGRNSPLAGASGFLGTTTGAIGPPLALIMRSMSPERVGTLS
VPIVGSIALITACALGMWHLQALVYAPAVIAGYFLSGVWKNLNRLLIFLGS
VTIISGSGWITQAGIF"
7592..8149
/notes="CE1188"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17998.1"
/db_xref="GI:23493027"
/translation="MGRDGFPAERPPADPRGRTIKMKWTKGAIJAAAAALLLV
GGGTWMSWDEVTQSDQIAAGNALSLADQWTLNGQDVNNVAVKVPDGLV
FTQIISVETSGDMLSATLTPGSIITGATADAADALATELNDITIDIGELPISITG
SGNTVTEGGOVSGSVSALSIKQAA"
complement(8190..9443)
/notes="CE1189, similar to AF024666-23|AAG03374.1| percent
identity: 84 in 408 aa"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC17999.1"
/db_xref="GI:23493028"
/translation="MAAGPYSIDPTTLYDELLAASPDLMREMQFINOILSTQADQ
VCGADYATVQDRVNTNGRHRDTRVGTIVAVPKLRTGSPFMDLLETRER
ALTTTCTCYLKGSTVRMDLVATLGNLSQVSDMAKDLQWVESPFRPLDTG
PVLVSCDALMKVRGGRVWTKSVLLATGVNAEGRYELLQWVATSVASVWTFGR
DLKAGLEVLVTSDAHLGIOHAIGVEIPNASQRCRTHSKNLYGMWSKYOWTFLS
AMPHIIFQPDPAQSVQAREVAFCEQKPPDVADYLEALDELAFNAFPAKSVTKV
WSNPTERINREIRRTDVVGFIPNRDVAIVRLVGVLAERQHDWILQQRVMSLSLEQ
TRAMTMANIISDDITSEIIRKQVA"
complement(11439..15143)
/gene="odha"
complement(11439..15143)

Query Match      34.5%; Score 818.6; DB 1; Length 308650;
Best Local Similarity 68.5%; Pred. No. 7.8e-190;
Matches 1182; Conservative 0; Mismatches 524; Indels 19; Gaps 3;

QY      650  TAACCATCAACATCAGTTTGATGCCCAATGCCGTCATCAACATCGCCAGCAGCAGTTG 709
DB      190920  TAACCATCAGATCAGCTTACGGCCATCCGGTGAGCACCACCGGCCACACCTATGTTG 190979

QY      710  ATCCAGCGCCACACCTTGGGGCTGGAACGGGGCTGACATGCTGTCGGCGGAACCC 769
DB      190980  ATCCAGCGCCAGACCGCGGGCTGGAACGGGGGACCGCAGTGGCGCGCGCGTAAACG 191039

QY      770  ACCAGGGGAACAGATCAGGCTTGCAGCGAAGCGCCAGCGGGGAAATCCACCGTCG 829
DB      191040  ACCAGTGGGAACAGACAGGCTTGGCGGCGAGGACCGCGAGCGAAGATCCACCGACG 191099

QY      830  GTGTGCGCGTATTCGCGCGCGACCGCGCGGATAAAACAACCGGTCCAAATACGATC 889
DB      191100  GTCTCCCGGTACTGGGCTCGACACACCGATGAAGACGAAGGATCCAGTAGGCATTG 191159

QY      890  GGGTTCAACAGGTGACGATGCGATCAACATGGCTTTTACCCTAAACCCGCTGCTTA 949
DB      191160  GGATTGAGCCAGTCAACATGGCCATGAGCATGGGCTGACCCAGACCTCCCGGTG 191219

QY      950  TCAGCGCTCACCTCCACCGCACCGGTGCGGTGCGTGTGAGTGGCCACCGCGCAACCG 1009
DB      191220  CCTGATGTGATCGGAGCGCGGGTCTGTTGTTGTCGTACGCGCCCGCGGAGCGGAC 191279

QY      1010  AAGAGCGTGTGATCGGCGACGGTGTGTTCTGTTTCTTCAATGATGTGTGGCGTTCC 1069

```

```

DB      191280  GTTGCGC---CAACGGGTTTGGGAATGCTC-----GACAAAGTTTACCTCGGTG 191324
QY      1070  TTGTTTGTATGCGGTCTTTTCGCTCCATGACGCGCAAAACATPAACAGGTAGGATGCA 1129
DB      191325  CGGGCGGCGAGGGGTGCGCGCGCCCATCACCGCAACACACAGCAGGTAGGGGATGCG 191384
QY      1130  CCCAGCGCATTAATATCGAGCAGCATCGCGCGGCATTTGGACAAAGATCAACGCCCAAG 1189
DB      191385  CACACGCGCAGGATGTCGAATATGATCGGGCGGTGTGCGAGATCGGCCGACCCCGAG 191444
QY      1190  GTCCGCGGATGAACAAAAAGAGCTCAGAAATTAACACACACGAGAAACCGCAATGAT 1249
DB      191445  GTCCGCGGATGAACAAAGACCACTCGGACGACGACGACGACGATGATCAGCGCGGTG 191504
QY      1250  CCTTCCGCGTTAATTCCTTTTAATCAACAGTACATCTTGGGTTCGATGAGCAGTAA 1309
DB      191505  CCCTCGGTTTGTATGCTCTGTTTGTATCACAGGACATTTCTGTGGGCGCATGGCCAC 191564
QY      1310  AGACTGCGCCCAAAAGCAGACCTGTAATGAAGATTTCCATGATCACCATCGTACCTAT 1369
DB      191565  AGACTGCGTCCCAACAAACCGGTGACAAAGATTTCCATGTCGCGCATGTCGCGCAC 191624
QY      1370  GGAAGTACTTAAGTAAATGATTTGTTCTTAACATGTTTAAATAGCTTCATGACCCC 1429
DB      191625  AAACCATTAAGTAAATGATGAGACTGAACATGCTTTAGAAITGCTTCATGACCCG 191684
QY      1430  ATTCACCTGACACTTTGCTCTCAATCATGATGAAGCAGCTTCGAAGGCGCTCTCTTA 1489
DB      191685  ATTCACCTGACACCTCTCTGACCATCATCGACGAGGCGAGCTTCGAGAACCTCTCTC 191744
QY      1490  GCCCTTTCCATTTCCCTCGGGCGGTGAGTCAGCGCTTAAAGCTCTCGAGATCACTG 1549
DB      191745  GCATCTGCGATTCCTCCCTTCGCGGTGACGCGCATCAAGCACTGAGAAATCCGTC 191804
QY      1550  GGTGAGTGTGTGATTCGCGCAACCAACCGGCAAGCAACCGAGCGGCTCAAGTCTTT 1609
DB      191805  GGTGCGGTACTGTGTGCGCACACCGCGCGGTGTCACCGAGCGCGGAGGTGCTG 191864
QY      1610  GTGCAAGCGCGGGAATGTTGCTCTCAGCAGCAAACTAAAGCGCACTATCTGGA 1669
DB      191865  GTCCAGCGCGCGCAAGATGGCGCTCTGTCAGGCGGAGACCCGTGAGCACTGCGGAA 191924
QY      1670  CGCTTGTGAAATCCGTTAACCATCGCCATCAAGCGAGATTCGCTATCCATGCTTT 1729
DB      191925  CGGCTTGACAGATTCCTCCCTGACCTGCTATCAAGCGGATTCCTGTCACCTGGTTC 191984
QY      1730  CTCTCCGTTTCAACGAGGTAGTCTTGGGGTGGAGCAAGCTCAGCTGCGCTTGGAA 1789
DB      191985  CCGCCCGTGTTCGCGAGGTTCGACATTTGGGGTGGCGTCACTCCCTGACGCTGCTG 192044
QY      1790  GATGAAGCGCACATTTATCTCTGCTGCGCGTGGAGATGTTTATGAGCGGTAAACCG 1849
DB      192045  GACGAGGCCACACCTCTGCTGTCGCGCGGTTCGCTGTCGCGGCGCGTGAACCGC 192104
QY      1850  GAAGCTAATCCCGTGGCGGATGTGAAGTAGTAGAACTTTGGAACCATCGCGCATTTG 1909
DB      192105  GAGCGGACCCCGTGTGCGGGTGTGAGGTCTTACGCTCGCGGTCTGAGACACCTGG 192164
QY      1910  ATTGCAACCCCTCATTTGCGGAGATGCTACATGTTGTGTGGAATACTAGATTGGG 1969
DB      192165  GTGGCCACCCCGGAACTGAGGCGCAGGTACACCGTGTGATGGCCAAACCCGACTGG 192224
QY      1970  ATGCGCGTCTTACGCTTCGCTCCCAAGATGCTGCTTCAAGACCGGTGACCTTGGAG 2029
DB      192225  ATGCGGTTACTGCGTTTTCGCGCGCGATGATGCTGTGTCGAGACCGGATCTCGAG 192284
QY      2030  GTGATGCTCTGTTGGGGCGCAGCGCGGTATTCATGTTCCCTCGCGGAGGTGTTGT 2089
DB      192285  GTGACGAGCGCTGCGCACGCGCGGTGTCTGCGTGTGCGGTCTGCGGAGGTTTGG 192344
QY      2090  GAGCAATTCGCGGAGGCTTGTGCGGACTTCTTCCCGMAACCCAAAGTGTCTCCATG 2149
DB      192345  GAGCGGTGCGCATCGCGCTGCGGCTGCGGCTGCTGCTCCCGAGGCGCGCGCGATG 192404

```

/translation="MRDMEIFVTGLIGASLLLAIGPQNVLVKQIKREGITAVIV
CLISDVFLTGLTGLVGLISDTAPILDLIRWCGIAYLLMAFAVMAARDLRARTEVTF
VHSPEVPAASAGSGVTKORPLRITTSQTRQVWRPMLMVLVTLWLPNPAIDAFV
FIGVGAQYGTGRMIFPAAGFAASLWPLVGYGAALSPPLSSPRVWRINIGAV
VLTLGVLKLLMG"

Query Match 33.6%; Score 798.8; DB 1; Length 1771;
Best Local Similarity 68.4%; Pred. No. 3.3e-185;
Matches 1158; Conservative 0; Mismatches 517; Indels 19; Gaps 3;

650 TAACCCATCAACATCAGTTTGTATGCGCAATCGGTGTATCACTGCGCAGACGAGTTG 709
1675 TAACCCATCAGGATCAGCTTCACGCGCAATCGGTGTAGCAGCAGCAGCAGCTATGTTG 1616

710 ATCCAGGCGCACCTTGGGGCTGGACAGCGCGGTGCAATGCTGCTGGCGCGAAACCC 769
1615 ATCCAGGCGCACCTTGGGGCTGGACAGCGCGGTGCAATGCTGCTGGCGCGCGGTAAACCG 1556

770 ACCAGCGGGAACCAAGATCAGGCTTGGCGGAAACCGCGCAGCGCGGAAATCCACGCTCG 829
1555 ACCAGTGGGAACCAAGATCAGGCTTGGCGGAAACCGCGCAGCGGAAATCCACGCGCG 1496

830 GTGTGCGCGTATGTGGCGCGCAGCGCGCGGTGCAATGCTGCTGGCGCGCGGTAAACCG 889
1495 GTCTCCCGCTACTGGGCTTCCGACACCAACCGATGAAGCAGGCGATCCAGGTAGGCAATTG 1436

890 GGGTTCAACACCGGTCAGCAGATGTCATCAACATGGGCTTACCAAAACCGCGCTGCTTA 949
1435 GGATTTGAGCGAGGTGAGCAATGCGCATGAGCATGAGCGCTGACCAAGCTGCGGGTG 1376

950 TCGACGCTCACCTCCACCGCGCAGCGGTTGGCGGTGTGCTGCTGCGCGCGCGCGCGCGAC 1009
1375 CCTGATGTATCGGAGCGCGGGTGTGTTGTTGCTGCTGCGCGCGCGCGCGCGCGAC 1316

1010 AAGGCGTGTATCGGCGCAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069
1315 GCTGCGCG---CAACGGGTTGCGAATGCTC-----GACAAAGGTTACCTCGGTG 1271

1070 TTGTTTGTATCGGCTCTTTTGGTGCATGAGCGGCAACCAACCAAGGTAAGCGATGCCA 1129
1270 CGGCGCGCAGGCGGTGCGCGCGCGCATCACCAGGAAACCAAGGTAAGCGATGCCG 1211

1130 CCCAGCGCATTAATCGAGCAGATCGCGCGCGCATTGGACAAAGATCAACGCCCGAG 1189
1210 CACCAGCGCGGATGTGAGAAATGATCGGCGCGGTGTCGAGATCAGCGCGCGCGCGAGG 1151

1190 GTGCGCGCATGACAAAGAGAGTCAAGAAATTAACACAGAGAGAGACCGCAATGAGT 1249
1150 GTGCGCGCGGTGACAGGACCGCGCGCGCATGATGATGATGATGATGATGATGATGATG 1091

1250 CTTTTCGCGCTTAATTCCTTTTAAATCACAGTACATTTCTGCGGTCCGATGGACAGTAA 1309
1090 CCTTCGCGGTGATGCTGCTGTTGATCACAGGACATTTCTGTCGCGCATGAGCGCGAGC 1031

1310 AGACTGCG 1369
1030 AGACTGCG 971

1370 GGAAGTACTTAAAGTAAATGATGCTTCTTAACATGTTTAAATATAGTCTTCAAGCGCG 1429
970 AAACCAACATTAAAGTAAATGATGAGACATGAGTAAATGATGATGATGATGATGATGAT 911

1430 ATTCAACTGACACACTTCTCTCAATCATTTGATGAGCGAGCTTTCGAGGCGCGCTCTTA 1489
910 ATCCACTTGGACACCTCTCTCAACCATCATCAGAGGAGCGAGCTTCGAGAAACGCTCTCT 851

1490 GCGCTTTTCCATTTTCCCGCTCGGCGGTGAGTCAAGCGGTAAAGCTCTCGAGCATCACTG 1549
850 GCACTGTGATCTCTCCCGCTCGGCGGTGAGTCAAGCGGTCAAGCGCATCGAGAAATCCGTC 791

1550 GGTGAGTGTGTGATTCGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1609

12150 CTAAGAGCAGAGAGTGTATCTCTCGATGAGTACCATCATGACACCGCATGTATGG 2209
192405 CTGGCTCGCGGGATGTGTGAGTGTGATGAGAGGGTGTGACACCGCGCTGTATGG 192464

2210 CACAGATGCGCGCTGGAATCTAGATCTAGCTAGACTCACAGACGCGGTGTATGCA 2269
192465 CACAGATGCGCGCTGGAATCTAGATCTAGCTAGACTCACAGACGCGGTGTATGCA 192524

2270 GCAATCGAGGATGTGGCGCTTGTATCTCTGAAAGGTTGAGGGTTTCACTTCTTC 2329
192525 GCGCGCGCGGTCTGTAG---AACCTAACTGGAACGCTTCCGCGGTCTTTCACCTCTTC 192580

2330 GCGCGCAGGATTTGGCGCAGGAGAGTAACTTCTGCAATG 2374
192581 ACCGAGGATATCGGCGCGCGCGAGGTTGACACCTCGCGGAGG 192625

RESULT 12
AB083133/c 1771 bp DNA linear BCT 06-APR-2002
DEFINITION Corynebacterium efficiens lysG, lysE genes for Lysine export transcriptional regulatory protein, Lysine exporter protein, complete cds.
ACCESSION AB083133
VERSION AB083133.1 GI:20065738
KEYWORDS Corynebacterium efficiens
SOURCE Corynebacterium efficiens
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1 Itaya, H., Kimura, E., Kawahara, Y. and Sugimoto, S. lysG, lysE of Corynebacterium efficiens Published Only in Database (2002) 2 (bases 1 to 1771)
AUTHORS Itaya, H., Kimura, E., Kawahara, Y. and Sugimoto, S.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) Hiroshi Itaya, AJINOMOTO CO., INC, Fermentation & Biotechnology Laboratories, 1-1, Suzuki-Cho, Kawasaki-Ku, Kawasaki, Kanagawa 210-8681, Japan (E-mail:hiroshi.itaya@ajinomoto.com, Tel:81-44-244-7123(ex.4146), Fax:81-44-222-0129)

FEATURES
Location/Qualifiers
1..1771
/organism="Corynebacterium efficiens"
/mol_type="genomic DNA"
/db_xref="taxon:152794"
complement(47..919)
/gene="lysG"
complement(47..919)
/gene="lysE"
/codon_start=1
/transl_table=11
/product="Lysine export transcriptional regulatory protein"
/protein_id="BAB88826.1"
/db_xref="GI:20065738"

translation="MNPILHDLTLTIIDEGSPENASLALISPSNVSRIKALEKSVG
RVLVSRTPQVATGAEVLVQAKRMALLOAETREQLAERLDEIPLTVAINADSLSTW
FPVFAVHAGVATLTLRVEDAHTLSLRGSLVATREADPVAGCEVLRLGVR
HLPVATPELRATVVDQPDWVMPVLRFGPNDVLQDRLEGRVDGAVARRVSVPS
ASGCEAVRLGLGWLPEAQAPMLAAGDVQVLDKVDVTPLYWQWRLESLRLARL
TDAVDARAGLR"

981..1676
/gene="lysE"
981..1676
/gene="lysE"
/codon_start=1
/transl_table=11
/product="Lysine exporter protein"
/protein_id="BAB88827.1"
/db_xref="GI:20065740"

Db 790 GGTGGGTACTGTGTGCGCACCCAGCCGCGTGTGCAACGAGCCGCGAGGTGCTG 731
 QY 1610 GTGCAAGCAGCGCGAAATGTGTGCTGCAAGCAGAACTAAAGCGCAACTATCTCGA 1669
 Db 730 GTCCAGCGCGCGCGAGATGCGCTGTGTCAGCGCGAGCCGTCGAGCAACTGCGGAA 671
 QY 1670 CGCTTGTCTGAATCCCGTTAAACCATGCGCAAGATGCTATCCATCGTGT 1729
 Db 670 CGCTTGTGACGAGATCCCGTGTATCAACGCGGATCCCTGTCCACCTGGTTC 611
 QY 1730 CTTCCGCTGTTCAAGCAGGTAGCTTCTGGGCTGGAGCAACGCTCAGCTGCGCTGGAA 1789
 Db 610 CCGCCCGTGTGCGCGAGGTGCAATTTGGGCTGCGTCACTGACGCTGCGTGGAG 551
 QY 1790 GATGAAGCGCACCAATTTATCTTGTGCGCGGTGGAGATGTTTTAGAGCGGTAAACCGT 1849
 Db 550 GACGAGCGCCACCCCTGTCTCTGTTGCGCGCGGTTCGCTGCGCGCGGTGACCCGC 491
 QY 1850 GAAGCTATCCGCTGGCGGATGTGAAGTAGTAGAATTTGGAACCATGCGCCACTTGGCC 1909
 Db 490 GAGCGCGACCCCGTGTGCGCGGTGTGAGGTCTTACGCTGCGGCTCATGACACCTGCCG 431
 QY 1910 ATTGAACCCCTCAATTCGCGGATGCTACATGCTTATGATGGGAAACTAGATTGGCTGCG 1969
 Db 430 GTGGCCACCCCGGAACTGAGGCGACGCTACACCTGATGGCCACCGACTGGGTGCGC 371
 QY 1970 ATGCCCGTCTTACGCTTGGTCCCAAGATGTCTTACAGCGGTGACCTGCGCGCGC 2029
 Db 370 ATGCCCGTACTGCGTTCGCGCGGATGATGTGTGCGAGGACCGCGATCTCGAGGCGCG 311
 QY 2030 GTGATGCTCTGCGCGCGCGGCTGATCTTCTGCGCGCGGAAAGTGTGCT 2089
 Db 310 GTGCGAGGAGCGCTGCCAGCGCGCTGCTCGCTGCTGCGCGGCTTTCGCG 251
 QY 2090 GAGCAATTCGCGCGAGCGCTTGTGTGGGACTTCTCCGAAACCCAAAGCTGCTCCCATG 2149
 Db 250 GAGCGGTGCGACTCGGCTGCGGCTGCGGCTGCTGCGGCGCGCGCGCGCGATG 191
 QY 2150 CTAAGCAGGAGAGTATCTCTGCTGATGATACCCATTCGACACCGCATCTATG 2209
 Db 190 CTGCTGCTCGCGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131
 QY 2210 CAAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2269
 Db 130 CAAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 71
 QY 2270 GCAATTCAGGAGATGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2329
 Db 70 GCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 15
 QY 2330 GCGCGCGGAAATG 2343
 Db 14 ACCGAGGCTATCG 1

RESULT_13

EX248357 349535 bp DNA linear BCT 06-NOV-2003
 LOCUS Corynebacterium diphtheriae gravis NC131123, complete genome;
 DEFINITION segment 4/8.
 ACCESSION EX248357 BX248353
 VERSION EX248357.1 GI:38199912
 KEYWORDS complete genome.
 SOURCE Corynebacterium diphtheriae
 ORGANISM Corynebacterium diphtheriae
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 1 (bases 1 to 349535)
 Cerdano-Tarraga A.M., Efstathiou A., Dover L.G., Holden M.T.G.,
 Pallen M., Bentley S.D., Besra G.S., Churcher C., James K.D., De
 Zoysa A., Chillingworth T., Cronin A., Dowd L., Felwell T.,
 Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
 Rabinowitz E., Rutherford K., Thomson N.R., Unwin L.,

TITLE Whitehead S. and Barrell B.G. Parkhill, J.
 JOURNAL The complete genome sequence and analysis of Corynebacterium
 PUBLISHED diphtheriae NC131123
 REFERENCE Nucleic Acids Res. 31 (22), 6516-6523 (2003)
 14602910
 2 (bases 1 to 349535)
 AUTHORS Cerdano-Tarraga A.M.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-2003) Cerdano-Tarraga A.M., submitted on behalf
 of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust
 Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:
 amct@sanger.ac.uk

FEATURES
 Location/Qualifiers
 1..349535
 /organism="Corynebacterium diphtheriae"
 /mol_type="genomic DNA"
 /strain="NC131123"
 /db_xref="taxon:1717"
 /note="biotype gravis"
 76..402
 /locus_tag="DIP1058"
 76..402
 /locus_tag="DIP1058"
 /note="Similar to Mycobacterium tuberculosis hypothetical
 11.0 kDa protein Rv1898 or MT1949 or MTCV180.20c
 SW:Y198 MYCTU (007734) (102 aa) fasta scores: E():
 5.8e-16, 56.43% id in 101 aa, and to Aquifex aeolicus
 hypothetical 17.9 kDa protein AQ.2067 TR:067847
 (EMBL:AB000771) (157 aa) fasta scores: E(): 4.7e-08,
 40.19% id in 102 aa"
 /codon_start=1
 /transl_table=11
 /product="Conserved hypothetical protein"
 /protein_id="CAB49578.1"
 /db_xref="GI:38199913"
 /translation="MLIDMLAFVAPPTETNNKAMADVAEIVRVRESGLPNETN
 AMFLIEGSEWVAVVVKATDAVLAVSPRAVLKADIRPGYTGLOQKVASVERVL
 AGDSSES"
 88..372
 /locus_tag="DIP1058"
 /note="HMPfam hit to PF01910, Protein of unknown function
 DUF77"
 complement(405..4762)
 /note="Putative iron uptake system. Low GC content
 (58.02%)"
 complement(405..1226)
 /locus_tag="DIP1059"
 complement(405..1226)
 /locus_tag="DIP1059"
 /note="Similar to Streptomyces coelicolor putative
 iron-siderophore uptake system ATP-binding component
 SC151.25C TR:09S215 (EMBL:AL109848) (301 aa) fasta scores:
 E(): 1.9e-42, 50.58% id in 255 aa, and to Escherichia coli
 ferric enterobactin transport ATP-binding protein fepC
 b0588 SW:PEPC.BCOLI (P23878) (271 aa) fasta scores: E():
 1.1e-41, 49.61% id in 262 aa"
 /codon_start=1
 /transl_table=11
 /product="Putative iron-siderophore uptake system
 ATP-binding component"
 /protein_id="CAE49580.1"
 /db_xref="GI:38199914"
 /translation="WTSTPILKTSVTLWDKHTVSTDLTVBPGQFTAIIGPNC
 GKSTLLKSFARILQPHQTCVTLGNPLDLSLHTKVIQRIALLPQALAPADITVEELV
 GRGFPYQTLKQWTDQAQVDSALLATGAQLHRVTELSGQQRQVWLMVLAQ
 NTPVVLIDEPYTLIDVAHQQLLEALRLQQLGRTIVTVHDLQQAQVADHLIYVK
 GGAIVAGAPSSITPQLISLSEFIDNVHTIDHVVIVKTPQPDVSVLENLQ"
 complement(564..1133)
 /locus_tag="DIP1059"
 /note="HMMsmart hit to SM00382, ATPases associated with a
 variety of cellular activities"
 complement(567..1130)
 /locus_tag="DIP1059"

misc_feature
 complement(405..4762)
 /note="Putative iron uptake system. Low GC content
 (58.02%)"
 misc_feature
 complement(405..1226)
 /locus_tag="DIP1059"
 misc_feature
 complement(405..1226)
 /locus_tag="DIP1059"

gene
 complement(405..1226)
 /locus_tag="DIP1059"

CDS
 complement(405..1226)
 /locus_tag="DIP1059"

misc_feature
 complement(564..1133)
 /locus_tag="DIP1059"

misc_feature
 complement(567..1130)
 /locus_tag="DIP1059"


```

/misc_feature
/note="HMPfam hit to PF00005, ABC transporter"
complement(579..797)
/locus_tag="DIP1059"
/note="ProfileScan hit to PS0100, 2nd half motif for
nucleotide binding, associated with P-loop."
complement(753..797)
/misc_feature
/locus_tag="DIP1059"
/note="ScanRegExp hit to PS00211, ABC transporters family
signature."
complement(1047..1124)
/misc_feature
/locus_tag="DIP1059"
/note="ProfileScan hit to PS0101, P-loop nucleotide
binding motif (does not find all)."
complement(1086..1109)
/misc_feature
/locus_tag="DIP1059"
/note="ScanRegExp hit to PS00017, ATP/GTP-binding site
motif A (P-loop)."
complement(1223..2182)
/misc_feature
/locus_tag="DIP1060"
complement(1223..2182)
/misc_feature
/locus_tag="DIP1060"
/note="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system transmembrane component
sc151.26C TR:Q9S214 (EMBL:AL109848) (375 aa) fasta scores:
E(): 5e-40, 40.74% id in 324 aa, and to Escherichia coli
ferric enterobactin transport system permease protein fepg
fep9 or B0589 SW:FEPC_ECOLI (P23877) (330 aa) fasta
scores: E(): 1.8e-45, 44.09% id in 322 aa"
/codon_start=1
/transl_table=11
/misc_feature
/product="Putative iron-siderophore uptake system
transmembrane component"
/protein_id="CAE49582.1"
/db_xref="GI:38199915"
/translation="MRALVLSLLVTAAYSLVPGAGLSTWQLTSETPLATVVM
QRLPRVTVGLVGAALAGSLFQSLTRNPLGSPDIFGSTGAYTVIAFLGWSG
FGATVIGALIGLVAVVIALSVTRIDGLRIILVGLISAMLSALNRMLITRGELD
TALSASWAGSLGRLVTAIPACCVLAITITPLRFLDLVSLGDLVAGLGLR
LQLKLLLVGVFLVAVTAVAGPIAFVALASPHARLTSRTPLVETSVIGALL
LIVADLIGRLRFQLFVGLVTVTIGMYLLNLIAKSKENS"
complement(2108..2182)
/misc_feature
/locus_tag="DIP1060"
/note="Signal peptide predicted for DIP1060 by SignalP 2.0
HMM (Signal peptide probability 0.966) with cleavage site
probability 0.180 between residues 25 and 26;
signal-peptide site"
complement(1244..2116)
/misc_feature
/locus_tag="DIP1060"
/note="HMPfam hit to PF01032, FecCD transport family"
order(complement(2105..2161), complement(1979..2044),
complement(1883..1939), complement(1805..1870),
complement(1718..1783), complement(1607..1657),
complement(1445..1510), complement(1319..1384),
complement(1247..1303))
/misc_feature
/locus_tag="DIP1060"
/note="Tg probable transmembrane helices predicted for
DIP1060 by TMHMM2.0"
complement(1250..1564)
/misc_feature
/locus_tag="DIP1060"
/note="BlastProDom hit to PD001557, PD001557"
complement(2184..3191)
/misc_feature
/locus_tag="DIP1061"
complement(2184..3191)
/misc_feature
/locus_tag="DIP1061"
/note="Similar to Streptomyces coelicolor putative
iron-siderophore uptake system transmembrane component
sc151.27C TR:Q9S213 (EMBL:AL109848) (348 aa) fasta scores:
E(): 8e-36, 36.53% id in 323 aa, and to Escherichia coli
ferric enterobactin transport system permease protein fepd
or B0590 SW:FEPC_ECOLI (P23876) (334 aa) fasta scores:
E(): 5.2e-40, 40.37% id in 322 aa"
/codon_start=1
/transl_table=11
/product="Putative iron-siderophore uptake system
transmembrane component"
/protein_id="CAE49584.1"
/db_xref="GI:38199916"
/translation="MLRQPRVRLPPLVITLVLIICTIIMSLGVGARNITSPDILA
ALRHGDPNHSIHWRIPTLIIAAGASISLAGVLIQALTRPLADTGVGINAG
AAFTVIGALIGLVAVVIALSVTRIDGLRIILVGLISAMLSALNRMLITRGELD
TALSASWAGSLGRLVTAIPACCVLAITITPLRFLDLVSLGDLVAGLGLR
LQLKLLLVGVFLVAVTAVAGPIAFVALASPHARLTSRTPLVETSVIGALL
LIVADLIGRLRFQLFVGLVTVTIGMYLLNLIAKSKENS"
complement(3090..3191)
/misc_feature
/locus_tag="DIP1061"
/note="Signal peptide predicted for DIP1061 by SignalP 2.0
HMM (Signal peptide probability 1.000) with cleavage site
probability 0.663 between residues 34 and 35;
signal-peptide site"
complement(2199..2525)
/misc_feature
/locus_tag="DIP1061"
/note="BlastProDom hit to PD001557, PD001557"
complement(2205..3095)
/misc_feature
/locus_tag="DIP1061"
/note="HMPfam hit to PF01032, FecCD transport family"
order(complement(3096..3161), complement(2934..2990),
complement(2838..2903), complement(2766..2831),
complement(2670..2735), complement(2544..2609),
complement(2394..2459), complement(2289..2354),
complement(2211..2276))
/misc_feature
/locus_tag="DIP1061"
/note="Tg probable transmembrane helices predicted for
DIP1061 by TMHMM2.0"
complement(3239..4282)
/misc_feature
/locus_tag="DIP1062"
complement(3239..4282)
/misc_feature
/locus_tag="DIP1062"
/note="Similar to Escherichia coli iron(III)
dicitrate-binding periplasmic protein precursor FecB or
B4290 SW:FEBC_ECOLI (P15028) (300 aa) fasta scores: E():
Query Match 30.5%; Score 724.4; DB 1; Length 349535;
Best Local Similarity 58.2%; Pred. No. 1.2e-166;
Matches 1351; Conservative 0; Mismatches 931; Indels 38; Gaps 3;
QY 1 AGATACCTCTTTGGAAGAACCAATGACGATTCGGTGACATTTGGCTCGGAAGGC 60
DB 35950 AGATACGCTTTGGAAGAAATGCATATGCTTTGCGAGACATCGGTCATCCGCAAGGC 36009
QY 61 TCTTTACGTGGTATTTCTTCAGTCCAGAGCTCACAGCGAGCGGCTGAGTTCAAT 120
DB 36010 TCCTATGTGGGAATTTCTCTTATGGACCTGAACCTAACCGAGAGATCTCGAATTTCT 36069
QY 121 GCGGAGGAGGCTGCCCGCTTCTGATTCATCAGCCAAAGCTATTCATCATTAATCGTTG 180
DB 36070 AGCTGTGAGGCTGCCCGCTGCTGATTCATCAACCGAGCTATCTATTTGAACCGCTG 36129
QY 181 GGTGGAGAACCGCGCGGATGACGGTGAGAACTTTGTCAGTCAGTCGCAACATGCTCT 240
DB 36130 GGTAGAAGAACCGGTGAAGATGGCGAGAACCTGTAGAGTCTGCGGGAATAACGAGCT 36189
QY 241 TGGCGTCATGCTTTCTCACACCTTCGCGCAGGCGCTCTCACGCAAAATATCTCGATGG 300
DB 36190 GGGTGTATTGATTTTGGCACTTTCACAGGGGCTTTTAACTGACAGTATCTCGAGGG 36249
QY 301 AATTCCAGAGGGTTCGCCCGCAGCGAGGTAAGTCCCTGTCTGAGGCGCATGTTGAACGT 360
DB 36250 CGTGCCAGTGGATTCACGGGACGACGAGGGAAGTCTTGGGCAAGAGATGCTCAACGC 36309
QY 361 GAACAATATTGATATGTCGCGAAGCTCAATGACATCGCCAGGACCGGCGGAGTCACCT 420
DB 36310 TAAAAATCTTGACATGGTGGCTGCACTGAATGATTTAGCGTTACGACGGCCAAAGCTT 36369
QY 421 TGGCGAGTGGCGCTTGCATGGGCTGTCGCGCAGCAAGGAGAGTACGCGCGGATACCGT 480

```

Db	36370	GGCACAGCTGGCTATTGTCGTGGGTCTCGTGGGAAACAAGGCGATTATATGGGGCCAAACAGGT	36429
Qy	481	GACCAAGTCATTGATTTGGTGGCTTCGTCAAGTTCAGCTGAGCTGGACAACACGCTTGATTCAC	540
Db	36430	AACCAAGTCGTTGATCGGAGCATCGTCGGTAGCGCAGTTGGATCAAAATCTGGAGACTCT	36489
Qy	541	CAACAACCTTGCAGTTTCTTGACCGCGAGTTTGGAGCGCATCGATGAGATTTTCCOACGAGC	600
Db	36490	TAATAATCTTGAGTTTCAGGCTTGAGAAACCAATATATTTGACACTGTTTGCAGAGCAGC	36549
Qy	601	CGCATCAACATTTTGGCGAAGGCGACCGATTCCTAAAACCGCGGAAAACCTAAC	653
Db	36550	TGGAAATTAATATTTGGCGGGTGTCTACTGCTCGAAGGTTTCACGATTAAGAGCGGGGTAC	36609
Qy	654	-----CCATCAACATCAGTTTGTATGGCCAATCGGCTCATCAACAT	694
Db	36610	CTTAAGAGCAATGAGCTTAGTGTGATGATGAGAGCTGGCGACATCATCATGATGATT	36669
Qy	695	GCCACGACAGCTTGATCCAGAGCGCACACCTTGGGGCTTGGACAGCGGGGCTGACATGCT	754
Db	36670	CCGATGCTTATGTTGATATATCGCCACACGCTGGCCGCGATAGCAGCTGAGAAAGCGT	36729
Qy	755	GCTGGCGCGAAACCCACACAGCGGGAAACAGATCAGCTTGGCCGGAACCGCGAGCGGCG	814
Db	36730	GTGGAGGTATATCCAAATAAACGGAACCACTGAGATCTGGCGCACTGGCCGAGGGCG	36789
Qy	815	AAATCCACGCTCGGTGTGGCGTATTGGCGCGGACCGCCGATTAACAACAACGCG	874
Db	36790	AAGACCCAGCGCCGCTGGAGCCGTTGTTGGCATTTCTCCACATGACAGAAACG	36849
Qy	875	TCCAAATACGCATTCGGGTTCAACACAGGTGACGACCATGGCATCAACATGGGCTTTACC	934
Db	36850	TCATATAAGCAGCTGGATTAAAGCCACGCTGAAGACCAACGCTGCAAGCACTGGTTAAC	36909
Qy	935	CAAAACCGCTGCTTATGAGCTCACCTTCCACCGCACCGCTGGCGGTGTCAGTGCC	994
Db	36910	CACGTGGGCTGC-----GCTGACTGGGCTGAGCTTTTGGCGTGTHTTTAGTATC	36960
Qy	995	ACCGCCGNAACCGCCCAAGGCGTGCATCGGCGACGTTGGTTCTGTTCTTCAATGATC	1054
Db	36961	ACTCCGATGATGCGGTAGCCACTGGTTCGTAGCCACCGGTTGCTTGTCTC--GACT	37017
Qy	1055	TGTGGCGCTTCCACCTTGTGTGTCATGGCGTCTTTCGTGCGCATGACGGCAACCATAC	1114
Db	37018	GGCAGCGCTTGGCGGTGACGTTTAATGTCTCTTAAACAAGTCAATTCGAATAGAGC	37077
Qy	1115	AGGTAAAGCATGCCACCCCGGCGATATATCGACGACGATCGCGGCGATTGGACAA	1174
Db	37078	AGTACGCTACACCGAGCCATTTTCAGTTACTTGTAGTCGATGGGGCTCGATCGACTAT	37137
Qy	1175	AGATCAACGCGCAAGGTGCGCGCGATGAACAAAGACGTCAGAAATTAACACACAGAG	1234
Db	37138	GCGCCGACACCGGAGTGCCTCCAAAATGAGATACGCTCAGAGATGTGCATACAAC	37197
Qy	1235	AGAACCGCAATGAGTCTTTCGGCGCTTAATTCCTGTTTATCACCATGATCTTCGGGT	1294
Db	37198	AGATGGGGAATTAGACCTTCGCGTTTAAATGCTTCAACGAATAATCAAGGCCATTTGTGC	37257
Qy	1295	CCGATGGAACAGTAAGACATGGCGCCCAAAAGACGCTGTGAATGAGATTTTCATGATC	1354
Db	37258	CCGATGCCCATGATCAGGACAGACCCATGAGGAATCCAGCGATAGCATCTACTACTA	37317
Qy	1355	ACCATCGTACACTATGAGTACTTAAGTAAATGATTTGTTCTTAAACATGTTTAAATAT	1414
Db	37318	AGCAATATCTCTCTGAAATGCTGAGTAAATTAACAGACTTTCAGGTTGGTTAGAAAT	37377
Qy	1415	AGCTTCATGAACCCCATTCACATGGAACATTTGCTCTCATCATTTGATGAAGGACGCTTC	1474
Db	37378	TGCTTCATGAATCCGCTTCATTTAGAAACCTTCTCGCAATAATTCGACGAGGACGCTTC	37437
Qy	1475	GAAAGCGCTCTCTAGCCCTTTCATTTCCCTCTGGCGGTGAGTCAGCGGCTTAAAGCT	1534
Db	37438	GAAAGTCGGGCTTTGATATAGGAATTTCTCTTACGACGATGATGAGGATTAAGCG	37497

QY	1535	CTCGAGCAATACGTGGGTCCGATGTGCTATGCGCACCCAAACGGCGCAAGCAACCGAA	1594
Db	37498	TTGGAGCGGGAGACGTGTGAGTGTAGTACGGCGTAGTCAACAGTCACTGCAACTGAT	37557
QY	1595	GCGGGTGAAGTCCCTTGTGCAACGACGCGGAAATGGTTGTGTGCAAGCAGAACTAAA	1654
Db	37558	GCGCGGAGGTACTGTGTGCAATCGGCACGAGTATGGCACTTCTGCAAGCCGAAACAAAC	37617
QY	1655	GCGCAACTATCTGAGCGCCTTGTGAAATCCCGTTAAACATCGCCATCAACGCAGATTG	1714
Db	37618	GCTCAGTTCGGCGGCAAGAAATGAGCGGAATCCATTTGCTGTAGCTGTCAACGAGATTCT	37677
QY	1715	CTATCCACATGGTTTCTCCCGTGTCAAAGAGTAGTCTTCTGGGTGTGAGCAACGATTC	1774
Db	37678	TTGGCCACATGGTTTTCGCGGGGTGATGCGAAGTGTGCGGAGATGGAATACGCCACATTG	37737
QY	1775	ACGTGCGCTTGGAAAGATGAAGCGCACACATTAATCTTGTCTGCGCGGTGAGCAACGTC	1834
Db	37738	CATATGGGTATCGAGCAGAGTCGCATTCGTTGTGTCTGTTGCGTCCGCGAGATGTCCT	37797
QY	1835	GGAGCGGTAAACCCGTGAAGCTAATTCCTGTGCGGGATGTGAAGTAGTAGAACTTGGAAAC	1894
Db	37798	GGCGCTGTAAACGGGGGAAAAGACACTGTATCTGTGTGCGACTCGATAGCGTTAGGTGCA	37857
QY	1895	ATGCGCACTTGGCCATTGGCAACCCCTCATTTGCGGATGCCCTACATGTGTGATGGGAAA	1954
Db	37858	TTCCGATATTTTGCAGTGGGCAGTCCGAAGCTGTGGATCGCTACCAATAGACGGTGCT	37917
QY	1955	CTAGATTGGGCTGCATGCCCGCTTTTAGCTTTCGGTTCGAAAGATGTGTCTCAAGACCGT	2014
Db	37918	GTTAATTTGGAGCGGATGCCAGTATTGGGATTTGGCCCTCGGGATGGGCTTCAAGACCGT	37977
QY	2015	GACCTGGAACGGCGCGTGCATGTCTGTGGGGCGACGGCGGTATCCATTGTCCCGTCG	2074
Db	37978	GATATTAGCGTCTGTCTGTGTGAAGTTCCGCGACGGCGGACGTGTTAATGAGATCCCTTCA	38037
QY	2075	GCGGAAGTTTGGTTAGGCAATTTCCGCGAGCGCTCTGGTTGGGGACTCTTTCGCGAAAC	2134
Db	38038	TCGGAGGCATTCATGGAAGCGGCCGTGTGGGATTTGGCTGGGCAATGCTCCCGCATTT	38097
QY	2135	CAAGCTGCTCCCATGCTAAAAGCAGGAGAAGTGAATCCCTCTGATGAGATACCCATTGAC	2194
Db	38098	CAAGCGTTACCGCTTTTGGATTTCTGTGTAGGTAGTGAATGCTGATAGCGAGGTGCAAGAA	38157
QY	2195	ACACCGATGTATTGGACAGATGGCGCTTGGAACTAGATCTCTACTAGACTCACAGAC	2254
Db	38158	GTTCGGCTGTATTGGCAGCGATGGCGGTTGGAAATCTCCAGCCCTAGAGCGGTTCACTCA	38217
QY	2255	GCGTCTGTTGATCCAGCAATCGAGGGAATTCGGCGCCTTAGT	2294
Db	38218	GCGGTAGTGGATGACGCCCTCGAGGGACTACATCCACTGT	38257

RESULT 14				
AX643028/c	AX643028	711 bp	DNA	linear PAT 24-FEB-2003
LOCUS	Sequence 7 from Patent EP1266966.			
DEFINITION	AX643028			
ACCESSION	AX643028.1	GI:28550158		
VERSION	Corynebacterium glutamicum			
KEYWORDS	Corynebacterium glutamicum			
SOURCE	Corynebacterium glutamicum			
ORGANISM	Bacteria; Actinobacteridia; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.			
REFERENCE	1	Gunji, Y. and Yasueda, H. Method for producing L-lysine or l-arginine by using methanol assimilating bacterium		
AUTHORS	Patient: EP 1266966-A 7 18-DEC-2002;			
TITLE	Ajinomoto Co., Inc. (JP)			
JOURNAL	Location/Qualifiers			
FEATURES	1..711			
SOURCE				

CDs

/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
1. 711
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD67811.1"
/db_xref="GI:28550159"
/db_xref="REMBL:CAD67811"
/translation="MWIWEIFITGILGASLLSIPQNVLVIKQIKREGLIAVLV
CLISDFPLAGTGLDLLSNAAPVLDIMWGGIAYLLFWFAWAAKADMTNVEAPQ
I1SEBTPVDPDTPLGSAVATDNRVSVSKORVWVPMMAVLVTLNPNAY
LDAPFVIGVGQAQVDTGRWIFAGAPFASLLWFLVPGFAGAAALSRLSPKVRWIN
VVAVVMTALAKMLMG"

ORIGIN

Query Match 29.9%; Score 711; DB 6; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.2e-163;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 CTAACCATCAACATCAGTTTGTATGGCAATGCGTCAATCAACATCGCACGACGCTT 708
DB 711 CTAACCATCAACATCAGTTTGTATGGCAATGCGTCAATCAACATCGCACGACGCTT 652
QY 709 GATCAGCGGCACACCTTGGGGCTGGACAGCGGGGTGACATGCTGCTGCGCGAAACC 768
DB 651 GATCAGCGGCACACCTTGGGGCTGGACAGCGGGGTGACATGCTGCTGCGCGAAACC 592
QY 769 CACACGCGGAAACAGATCAGGCTTTGCCGGAACCGCGCAGCGCGAAATCCACCGTCC 828
DB 591 CACACGCGGAAACAGATCAGGCTTTGCCGGAACCGCGCAGCGCGAAATCCACCGTCC 532
QY 829 GGTGTGCGCGTATTGCGCGCGAAGCGCGCGGATTAACAAACGGGTGCAATACGCAAT 888
DB 531 GGTGTGCGCGTATTGCGCGCGAAGCGCGCGGATTAACAAACGGGTGCAATACGCAAT 472
QY 889 CGGGTTCAACAGGTGACGAGTATGCGTCAATCAATGCGGCTTACCCAAACCGCTGCTT 948
DB 471 CGGGTTCAACAGGTGACGAGTATGCGTCAATCAATGCGGCTTACCCAAACCGCTGCTT 412
QY 949 ATCGAGCTCACTTCAACCGGACCGCGGTGCGGTGTCAGTGGCGACCGCGAACCGCC 1008
DB 411 ATCGAGCTCACTTCAACCGGACCGCGGTGCGGTGTCAGTGGCGACCGCGAACCGCC 352
QY 1009 CAAAGCGGTGTCAGCGGACCGGTGCGGTGTCAGTGGCGACCGCGAACCGCC 1068
DB 351 CAAAGCGGTGTCAGCGGACCGGTGCGGTGTCAGTGGCGACCGCGAACCGCC 292
QY 1069 CTGTTGTTGTCAGCGGTGTCAGTGGCGACCGGTGTCAGTGGCGACCGCGAACCGCC 1128
DB 291 CTGTTGTTGTCAGCGGTGTCAGTGGCGACCGGTGTCAGTGGCGACCGCGAACCGCC 232
QY 1129 ACCCCAGCGCATATATGACGACGATCGCGCGGATTTGGAACAAAGATCAACCGCCAA 1188
DB 231 ACCCCAGCGCATATATGACGACGATCGCGCGGATTTGGAACAAAGATCAACCGCCAA 172
QY 1189 GGTGCGCGGATGACAAAGAGGTGAGAAATTAACACACGAGAAACCGCAATGAG 1248
DB 171 GGTGCGCGGATGACAAAGAGGTGAGAAATTAACACACGAGAAACCGCAATGAG 112
QY 1249 TCCTTCGCGCTTAATCTCTGTTTAAATCACCAGTACATCTCGCGGTCCGATGGACAGTAA 1308
DB 111 TCCTTCGCGCTTAATCTCTGTTTAAATCACCAGTACATCTCGCGGTCCGATGGACAGTAA 52
QY 1309 AAGACTGGCCCCAAAAGCAGACCTGTAAATGAAGATTTCATGATCAACAT 1359
DB 51 AAGACTGGCCCCAAAAGCAGACCTGTAAATGAAGATTTCATGATCAACAT 1

RESULT 15
AX123539/c 708 bp DNA linear PAT 11-MAY-2001
LOCUS AX123539

DEFINITION Sequence 3455 from Patent EP1108790.
ACCESSION AX123539
VERSION AX123539.1 GI:14041027
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriidae; Corynebacterium.
1
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
Novel polynucleotides
Patent: EP 1108790-A 3455 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
LOCATION/Qualifiers
1. 708
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN

Query Match 29.8%; Score 708; DB 6; Length 708;
Best Local Similarity 100.0%; Pred. No. 6.5e-163;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 ACCCATCAACATCAGTTTGTATGGCAATGCGTCAATCAACATCGCACGACGCTTGT 711
DB 708 ACCCATCAACATCAGTTTGTATGGCAATGCGTCAATCAACATCGCACGACGCTTGT 649
QY 712 CGAGCGGCACACCTTGGGGCTGGACAGCGGGGTGACATGCTGCTGCGCGAAACCAC 771
DB 648 CGAGCGGCACACCTTGGGGCTGGACAGCGGGGTGACATGCTGCTGCGCGAAACCAC 589
QY 772 CAGCGGAAACAGATCAGGCTTGGCGCGAACCGCGCGGCGGAAATTCACCGTCCGT 831
DB 588 CAGCGGAAACAGATCAGGCTTGGCGCGAACCGCGCGGCGGAAATTCACCGTCCGT 529
QY 832 GTGCGCGTATTGCGCGCGACCGCGCGGATTAACAAACGGGTGCAATACGCAATCGG 891
DB 528 GTGCGCGTATTGCGCGCGACCGCGCGGATTAACAAACGGGTGCAATACGCAATCGG 469
QY 892 GTTCAACAGGTGACGACGATGCGTCAATCAATGCGGCTTACCCAAACCGCTGCTTATC 951
DB 468 GTTCAACAGGTGACGACGATGCGTCAATCAATGCGGCTTACCCAAACCGCTGCTTATC 409
QY 952 GAGCTCACTTCAACCGCGACCGCGGTGCGGTGTCAGTGGCGACCGCGAACCGCCAA 1011
DB 408 GAGCTCACTTCAACCGCGACCGCGGTGCGGTGTCAGTGGCGACCGCGAACCGCCAA 349
QY 1012 AGGCGTGTCAATCGGCGACGCTGCTGCTTCTTCAATGATCTGCGGCTTCCACCTT 1071
DB 348 AGGCGTGTCAATCGGCGACGCTGCTGCTTCTTCAATGATCTGCGGCTTCCACCTT 289
QY 1072 GTTGTTCATGGCGCTTTCGCTGCGATGACGCGAAACCATACAGGTAGCGATGCGCAC 1131
DB 288 GTTGTTCATGGCGCTTTCGCTGCGATGACGCGAAACCATACAGGTAGCGATGCGCAC 229
QY 1132 CGAGCGCATATATGACGACGATCGCGCGGATTTGGAACAAAGATCAACCGCCAAAGT 1191
DB 228 CGAGCGCATATATGACGACGATCGCGCGGATTTGGAACAAAGATCAACCGCCAAAGT 169
QY 1192 GCGCGCGATGACAAAGAGAGCTGAGAAATTAACACACGAGGAGGACCGCAATGAGTCC 1251
DB 168 GCGCGCGATGACAAAGAGAGCTGAGAAATTAACACACGAGGAGGACCGCAATGAGTCC 109
QY 1252 TTCCGCGCTTAATCTCTGTTTAAATCACCAGTACATCTTGGCGGTCCGATGGACAGTAAAG 1311
DB 108 TTCCGCGCTTAATCTCTGTTTAAATCACCAGTACATCTTGGCGGTCCGATGGACAGTAAAG 49
QY 1312 ACTGCGCCCCAAAAGCAGACCTGTAAATGAAGATTTCATGATCAACAT 1359
DB 48 ACTGCGCCCCAAAAGCAGACCTGTAAATGAAGATTTCATGATCAACAT 1

Thu Mar 18 12:31:04 2004

us-09-105-117k-3.rge

Page 20

Search completed: March 15, 2004, 22:02:11
Job time : 9512 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 10:09:44 ; Search time 968.046 Seconds
(without alignments)
10418.122 Million cell updates/sec

Title: US-09-105-117K-3
Perfect score: 2374
Sequence: 1 agatactccttgaagaaa.....gtacacaccttcagcaaatgg 2374

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124093041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_25Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	2374	100.0	2374	2 AAT96816	Aat96816 DNA encod
C 2	2374	100.0	2374	9 ADB66196	Adh66196 DNA fragm
C 3	2374	100.0	349980	5 AAH68528	Aah68528 C glutami
C 4	993	41.8	993	4 AAF71777	Aaf71777 Coryneb
C 5	993	41.8	993	4 AAF71779	Aaf71779 Coryneb
C 6	870	36.6	870	5 AAH68421	Aah68421 C glutami
C 7	822	34.6	822	4 AAF71779	Aaf71779 Coryneb
C 8	822	34.6	822	4 AAH68421	Aah68421 C glutami
C 9	711	29.9	711	8 ACC80941	Acc80941 LysE prot
C 10	708	29.8	708	5 AAH68420	Aah68420 C glutami
C 11	698.6	29.4	1568	4 AAH45375	Aah45375 C. thermo
C 12	696.8	29.4	712	8 ACC80942	Acc80942 LysE24 pr
C 13	627	26.4	627	5 AAH68419	Aah68419 C glutami
C 14	485.2	20.4	993	4 AAF68077	Aaf68077 Coryneb
C 15	305.8	12.9	1095	7 ACA29651	ACA29651 Prokaryot
C 16	302.8	12.8	879	7 ACA29653	ACA29653 Prokaryot
C 17	132.6	5.6	1041	7 ACA25567	ACA25567 Prokaryot
C 18	123.8	5.2	897	7 ACA26879	ACA26879 Prokaryot
C 19	122.8	5.2	909	7 ACA38559	ACA38559 Prokaryot
C 20	122.8	5.2	15239	2 AAT33536	Aat33536 BCG delet
C 21	121.2	5.1	912	7 ACA40574	ACA40574 Prokaryot
C 22	121.2	5.1	110000	4 AAT99682_22	Continuation (23 o
C 23	121.2	5.1	110000	4 AAT99683_22	Continuation (23 o

24	116.4	4.9	1041	7	ACA24017	Prokaryot
25	114.8	4.8	1107	7	ACA26520	Prokaryot
26	108.6	4.6	1038	7	ACA32095	Prokaryot
C 27	108	4.5	1207	9	ADD13356	C. glutam
28	107	4.5	1041	7	ACA19150	Prokaryot
29	105.2	4.4	999	7	ACA15119	Prokaryot
30	105.2	4.4	999	7	ACA50894	Prokaryot
C 31	101	4.3	1200	4	AAF71729	Coryneb
C 32	100.8	4.2	5541	2	AAQ55755	Escherich
33	97.4	4.1	1035	7	ACA45509	Prokaryot
34	93.2	3.9	988	6	ABK72786	Bacillus
35	90.4	3.8	349980	6	ABQ81844	Bifidobac
36	82.6	3.5	894	4	AAH52558	E. coli D
37	82.6	3.5	894	7	ACA32622	Prokaryot
38	81.4	3.4	999	7	ACA24884	Prokaryot
39	79.8	3.4	909	7	ACA53702	Prokaryot
40	79.4	3.3	23128	4	AAH59552	Propionib
41	79.4	3.3	23128	7	ACF64481	Propionib
42	78.2	3.3	990	7	ACA35500	Prokaryot
C 43	78	3.3	1077	5	AAH66360	C glutami
44	77.6	3.3	903	7	ACF71729	Photobab
C 45	77.6	3.3	110000	7	ACF67367_49	Continuation (50 o

ALIGNMENTS

RESULT 1
AAT96816/c
ID AAT96816 standard; DNA; 2374 BP.
XX
AC AAT96816;
XX
DT 12-MAR-1998 (first entry)
XX
DE DNA encoding LysG, LysE and ORF3 from Corynebacterium glutamicum.
XX
LysG; LysE; ORF3; lysine transport; regulatory protein; export protein;
KW Microbial production; amino acid; animal feed additive; ds.
XX
OS Corynebacterium glutamicum.
XX
FH Key Location/Qualifiers
FT CDS complement(82..954)
FT /tag= a
FT /label= LysG
FT 1016..1726
FT /tag= b
FT /label= LysE
FT CDS complement(1723..2373)
FT /tag= c
FT /label= orf3
DEI9548222-A1.
26-JUN-1997.
22-DEC-1995; 95DE-01048222.
22-DEC-1995; 95DE-01048222.
(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
Vrijic M, Eggeling L, Sahn H;
WPI; 1997-333867/31.
P-PSDB; AAW37714, AAW37715, AAW37716.
Increasing microbial production of amino acids, especially lysine - by
improving export carrier activity or corresponding gene expression, also
new export and regulatory genes from Corynebacterium.
Claim 23 and 26; Page; 16pp; German.

XX This DNA, isolated from *Corynebacterium glutamicum*, contains the *lysG*,
 CC *lysE* and *ORF3* genes. *lysG* and *lysE* encode a lysine transport regulatory
 CC protein and an export protein, respectively. Microbial production of
 CC amino acids (A) is improved by increasing the export-carrier activity
 CC and/or the export gene expression in a microorganism that produces (A).
 CC The method is specifically used to increase production of lysine, used as
 CC an animal feed additive. Other (A) are variously useful as
 CC pharmaceuticals, condiments and intermediates for fine chemicals. This
 CC method increases the amount of (A) secreted into the culture medium.
 CC Export of (A) has been found to depend on a single gene. NB. This
 CC sequence has been created from the information given in table 2 of the
 CC specification
 XX
 SQ Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;
 Query Match 100.0%; Score 2374; DB 2; Length 2374;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGATACCTCTTGGAGAACCACTGTAGCAATTGGCGTACATTGTTGGCTCTGAAAGGC 60
 DB 2374 AGATACCTCTTGGAGAACCACTGTAGCAATTGGCGTACATTGTTGGCTCTGAAAGGC 2315
 QY 61 TCTTTACGTGGGTATTTCTTCTTACGGTCCAGAGCTCACAGCGGAGGCGGTGAGTTTCA 120
 DB 2314 TCTTTAGGTGGGTATTTCTTCTTACGGTCCAGAGCTCACAGCGGAGGCGGTGAGTTTCA 2255
 QY 121 GCGGAGGAGGGCTGCCCGCTTCTGATTCATCAGCCAGCTATTCCATCATTAATCGTTG 180
 DB 2254 GCGGAGGAGGGCTGCCCGCTTCTGATTCATCAGCCAGCTATTCCATCATTAATCGTTG 2195
 QY 181 GGTGGAGGAACCGGGCGATGACGGTGAGAACTTGTGTCAGTCAGTCGCCAACAAATGCT 240
 DB 2194 GGTGGAGGAACCGGGCGATGACGGTGAGAACTTGTGTCAGTCAGTCGCCAACAAATGCT 2135
 QY 241 TGGCGTCATTTGCTTCTCACCACTTGGCGAGGGCTGCTCACGGAACAAATATCTCGATGG 300
 DB 2134 TGGCGTCATTTGCTTCTCACCACTTGGCGAGGGCTGCTCACGGAACAAATATCTCGATGG 2075
 QY 301 AATTCAGAGGGTTCGCGCGCAGCGAGGGTAAGTCCCTGCTGAGGGCATGTTGAAGCT 360
 DB 2074 AATTCAGAGGGTTCGCGCGCAGCGAGGGTAAGTCCCTGCTGAGGGCATGTTGAAGCT 2015
 QY 361 GAACAATATTGATATGTCGCGCAAGCTCAATGACATCGCCAGGAACCGCGGCACTCACT 420
 DB 2014 GAACAATATTGATATGTCGCGCAAGCTCAATGACATCGCCAGGAACCGCGGCACTCACT 1955
 QY 421 TGGCGAGATGGCGCTTGCATGGGTGTCGCGGAGCAAGGAGATACGCGCGGATACCGT 480
 DB 1954 TGGCGAGATGGCGCTTGCATGGGTGTCGCGGAGCAAGGAGATACGCGCGGATACCGT 1895
 QY 481 GACCACTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 1894 GACCACTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1835
 QY 541 CAACAACCTTGGAGTTTCTGACCGCGAGTTGGAGCGCATCGATGAGATTTCCACGAGCG 600
 DB 1834 CAACAACCTTGGAGTTTCTGACCGCGAGTTGGAGCGCATCGATGAGATTTCCACGAGCG 1775
 QY 601 CGGATCAACATTTGGGGAAGGCGCACCGATTCGAAACCGCGGAAAACCTAACCCATCAA 660
 DB 1774 CGGATCAACATTTGGGGAAGGCGCACCGATTCGAAACCGCGGAAAACCTAACCCATCAA 1715
 QY 661 CATCAGTTTGTATGGCCATGCGGTGATCACTCACTGCAAGAGAGTTGATCCAGCGCA 720
 DB 1714 CATCAGTTTGTATGGCCATGCGGTGATCACTCACTGCAAGAGAGTTGATCCAGCGCA 1655
 QY 721 CACCTTGGGCGTGACAGCGGGGTGACAAATGCTGCTGCGCGGAAACCGCACCGCGGAA 780
 DB 1654 CACCTTGGGCGTGACAGCGGGGTGACAAATGCTGCTGCGCGGAAACCGCACCGCGGAA 1595
 QY 781 CCAGATCAGGCTTGGCGGAAACGCGCCAGCGGCGGAAATCCACCGTGGGTGTCGCGGTA 840

DB 1594 CCAGATCAGGCTTGGCGGAGCGCGCAGCGCGGAAATCCACCGTCCGGTGTGCGCGTA 1535
 QY 841 TTGGGCGCGAGCGCGCGGATTAACACAAAACGGTCCAAATAGCAATTCGGGTTCACCA 900
 DB 1534 TTGGGCGCGAGCGCGCGGATTAACACAAAACGGTCCAAATAGCAATTCGGGTTCACCA 1475
 QY 901 GGTTCAGCAGATTGTCATCAACATGAGGTGTTTACCACAAACCGCTCTTATCGAGCTCAC 960
 DB 1474 GGTTCAGCAGATTGTCATCAACATGAGGTGTTTACCACAAACCGCTCTTATCGAGCTCAC 1415
 QY 961 CTCACCCCGACCGGTGGCGGTGTCAGTGGCCACCGCGGAAACCGCCAAAGCGGTGTC 1020
 DB 1414 CTCACCCCGACCGGTGGCGGTGTCAGTGGCCACCGCGGAAACCGCCAAAGCGGTGTC 1355
 QY 1021 ATCGGCGACGGTGGTCTGCTTCTTCAATGATCTGTGGCGTTCACCTGTTGTTGTCAT 1080
 DB 1354 ATCGGCGACGGTGGTCTGCTTCTTCAATGATCTGTGGCGTTCACCTGTTGTTGTCAT 1295
 QY 1081 GGGTCTTCTTGGTCCATGACGCAACCAATACAGGTGAAGCGATGCCACCCAGCGCAT 1140
 DB 1294 GGGTCTTCTTGGTCCATGACGCAACCAATACAGGTGAAGCGATGCCACCCAGCGCAT 1235
 QY 1141 AATATGACGACGATCGCGCGGCAATTGGACAAAAGATCAACGCCAAGGTGCGCGGAT 1200
 DB 1234 AATATGACGACGATCGCGCGGCAATTGGACAAAAGATCAACGCCAAGGTGCGCGGAT 1175
 QY 1201 GAACAAAAGAGCGTCAGAAATTAACACACGAGAAAGAACCGCAATGAGTCTTTCGCGCTT 1260
 DB 1174 GAACAAAAGAGCGTCAGAAATTAACACACGAGAAAGAACCGCAATGAGTCTTTCGCGCTT 1115
 QY 1261 AATTCCTTCTTAACTACCACTACCTTTCGCGTCCGATGGAACAGTAAAGACATGCGCCC 1320
 DB 1114 AATTCCTTCTTAACTACCACTACCTTTCGCGTCCGATGGAACAGTAAAGACATGCGCCC 1055
 QY 1321 CAAAGCAGACCTGTATGAGATTTCCATGATCAGCTGATGATGATGATGATGATGATGAT 1380
 DB 1054 CAAAGCAGACCTGTATGAGATTTCCATGATCAGCTGATGATGATGATGATGATGATGAT 995
 QY 1381 AGTAAATCATTTGTTCTTAACTGTTTAAATATATAGTCTTCAAGACCCCATTTCACTGGA 1440
 DB 994 AGTAAATCATTTGTTCTTAACTGTTTAAATATATAGTCTTCAAGACCCCATTTCACTGGA 935
 QY 1441 CACTTTGCTCTCAATCATTTGATGAAGCAGCTTGAAGGCGCTCTCTAGCCCTTTCCAT 1500
 DB 934 CACTTTGCTCTCAATCATTTGATGAAGCAGCTTGAAGGCGCTCTCTAGCCCTTTCCAT 875
 QY 1501 TTCCCTCTCGCGGTGATGATGAGCGCTTAAAGCTCTCGAGCATCAAGTGGGTGAGTGT 1560
 DB 874 TTCCCTCTCGCGGTGATGATGAGCGCTTAAAGCTCTCGAGCATCAAGTGGGTGAGTGT 815
 QY 1561 GGTATCGCGCACCAACCGGCGAAAGCAACCGAGCGGGTGAAGTCTTGTGCAAGCAGC 1620
 DB 814 GGTATCGCGCACCAACCGGCGAAAGCAACCGAGCGGGTGAAGTCTTGTGCAAGCAGC 755
 QY 1621 GCGGAAATGTTGTTGTCGACGAGCAAACTTAAGCGCAACTATCTGAGCGCTTGTGGA 1680
 DB 754 GCGGAAATGTTGTTGTCGACGAGCAAACTTAAGCGCAACTATCTGAGCGCTTGTGGA 695
 QY 1681 AATCCGTTAAACCATCGCATCAACGAGATTCGCTATCCCATGTTTCTCCCGTGT 1740
 DB 694 AATCCGTTAAACCATCGCATCAACGAGATTCGCTATCCCATGTTTCTCCCGTGT 635
 QY 1741 CAACGAGTAGCTTCTTGGGTGGAGCAACCGCTCACGCTCGCTTGGAGATGAAGCGCA 1800
 DB 634 CAACGAGTAGCTTCTTGGGTGGAGCAACCGCTCACGCTCGCTTGGAGATGAAGCGCA 575
 QY 1801 CACATTTATCTTGTGCGCGGTGAGATGTTTGGAGCGGTAACCGGTGAGCTATTC 1860
 DB 574 CACATTTATCTTGTGCGCGGTGAGATGTTTGGAGCGGTAACCGGTGAGCTATTC 515
 QY 1861 GTCGCGGATGTAAGTAGTAGAATTGGAAACCATGCGCACTTGGGCGATTTGCAACCCC 1920

Db 514 CCGTGGCGGATGTGAAGTAGTAGAACTTGGAAACCATGGCCACCTTGGCCATTCGAAACCCC 455
 QY 1921 CTCATTGGGGATGCTACATGCTGTGATGGAACTAGATTGGGCTGGCATCCCGCTTT 1980
 Db 454 CTCATTGGGGATGCTACATGCTGTGATGGAACTAGATTGGGCTGGCATCCCGCTTT 395
 QY 1981 AGCTTTCGCTCCCAAGATGTCTTCAAGACCGTGACCTGGACGGCGGCTCGATGGTCC 2040
 Db 394 AGCTTTCGCTCCCAAGATGTCTTCAAGACCGTGACCTGGACGGCGGCTCGATGGTCC 335
 QY 2041 TGTGGGCGGACGGCGGTATCATATGCTCCGTCGCGGAGGTTTTGGTGGGCAATTCG 2100
 Db 334 TGTGGGCGGACGGCGGTATCATATGCTCCGTCGCGGAGGTTTTGGTGGGCAATTCG 275
 QY 2101 CCGAGGCTTGTGGGCACTTCTCCGGAACCCCAAGCTGCTCCCATCTTAAGCAGG 2160
 Db 274 CCGAGGCTTGTGGGCACTTCTCCGGAACCCCAAGCTGCTCCCATCTTAAGCAGG 215
 QY 2161 AGAAGTGATCTCTCGATGAGATACCATTTGACACACCGATGTATTGGCAACGATGGG 2220
 Db 214 AGAAGTGATCTCTCGATGAGATACCATTTGACACACCGATGTATTGGCAACGATGGG 155
 QY 2221 CCGAATCTAGATCTCTAGCTAGATCTACAGACCGCTGCTGATGAGCAATCGAGG 2280
 Db 154 CCGAATCTAGATCTCTAGCTAGATCTACAGACCGCTGCTGATGAGCAATCGAGG 95
 QY 2281 ATTGGGCGCTTACTTCTTGAAGGTTTCAGGGTTTTTCACCTTCTCGCCCGCAGGAA 2340
 Db 94 ATTGGGCGCTTACTTCTTGAAGGTTTCAGGGTTTTTCACCTTCTCGCCCGCAGGAA 35
 QY 2341 TTGGGCCAGCAGAGTAACACCTTCAGCAAAATGG 2374
 Db 34 TTGGGCCAGCAGAGTAACACCTTCAGCAAAATGG 1

RESULT 2
 ADB66196/c
 ID ADB66196 standard; DNA; 2374 BP.
 XX AC ADB66196;
 XX DT 04-DEC-2003 (first entry)
 XX DE DNA fragment containing C. glutamicum LysE and LysE genes.
 XX KW L-arginine production; coryneform bacteria; LysE; arginine repressor;
 XX KW argR; liver function promoting agent; amino acid infusion;
 XX KW amino acid pharmaceutical; LysE; ds.
 XX OS Corynebacterium glutamicum.
 XX FT Key Location/Qualifiers
 XX FT CDS 1025..1726
 XX FT /*tag= a
 XX FT /product= "protein encoded by LysE gene"
 XX PN US2003113899-A1.
 XX PD 19-JUN-2003.
 XX PF 17-JUL-2002; 2002US-00196232.
 XX PR 25-JUL-2001; 2001JP-00224586.
 XX PA (AJIN) AJINOMOTO CO INC.
 XX PI Yamauchi M, Ito H, Gunji Y, Yasueda H;
 XX DR WPI; 2003-708853/67.
 XX DR P-PSDB; ADB66197.
 XX FT A microorganism comprising enhanced expression of the lysE gene is useful
 for enhanced production of L-arginine.

XX Example 4; Page 26-27; 36pp; English.
 PS The present invention relates to a method for producing L-arginine in a
 CC microorganism (e.g. coryneform bacteria) that has L-arginine producing
 CC ability and has been modified for enhanced expression of the lysE gene.
 CC The microorganism is also modified so that an arginine repressor (argR)
 CC does not function normally. The method of the invention is useful for the
 CC enhanced production of L-arginine which is useful in liver function
 CC promoting agents, amino acid infusion and comprehensive amino acid
 CC pharmaceuticals. The present sequence represents a DNA fragment
 CC containing Corynebacterium glutamicum LysE and LysE genes. Note: The
 CC present sequence is given as SEQ ID No.24 in the Sequence Listing but is
 CC referred to as SEQ ID No.25 in the rest of the specification.
 XX Sequence 2374 BP; 526 A; 640 C; 648 G; 560 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 2374; DB 9; Length 2374;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGAATCTCTTTGGAAGAAACCATGTAGCATTTGGCTGACATTTGCTGCTGGAAGGC 60
 Db 2374 AGATATCTCTTTGGAAGAAACCATGTAGCATTTGGCTGACATTTGCTGCTGGAAGGC 2315
 QY 61 TCTTTACGTGGGTATTTCTTCTACGGTCCAGAGCTTCACAGCGAGCGGCTGAGTTTCA 120
 Db 2314 TCTTTACGTGGGTATTTCTTCTACGGTCCAGAGCTTCACAGCGAGCGGCTGAGTTTCA 2255
 QY 121 GCGGAGGAGGGTTCGCCCTTCTGATTCATCAGCAAGCTATTCCATCATTAATCGTTG 180
 Db 2254 GCGGAGGAGGGTTCGCCCTTCTGATTCATCAGCAAGCTATTCCATCATTAATCGTTG 2195
 QY 181 GGTGAGGAAACCGGCGGATGACGGTGGAGACTTTGTGACGTCACTGCCCAATATGTTCT 240
 Db 2194 GGTGAGGAAACCGGCGGATGACGGTGGAGACTTTGTGACGTCACTGCCCAATATGTTCT 2135
 QY 241 TGGCGTCAATTGCTTTCTCACCACTTGGCGAGGCGCTGCTCACGCAAAATATCTCGATGG 300
 Db 2134 TGGCGTCAATTGCTTTCTCACCACTTGGCGAGGCGCTGCTCACGCAAAATATCTCGATGG 2075
 QY 301 AATTCAGAGGGTTCCCGCGCCAGCGAGGTAACTCCCTGCTCGAGGGCATGTTGAACGT 360
 Db 2074 AATTCAGAGGGTTCCCGCGCCAGCGAGGTAACTCCCTGCTCGAGGGCATGTTGAACGT 2015
 QY 361 GACCAATATTGATGGTCCGAGCTCATGACATCCGCCAGCAAGCGGGGCACTCACT 420
 Db 2014 GACCAATATTGATGGTCCGAGCTCATGACATCCGCCAGCAAGCGGGGCACTCACT 1955
 QY 421 TGGCGAGATGGCGCTTGCATGGGTGCTGCGCGAGCAAGAGAGTACGGCGCGGATACCGT 480
 Db 1954 TGGCGAGATGGCGCTTGCATGGGTGCTGCGCGAGCAAGAGAGTACGGCGCGGATACCGT 1895
 QY 481 GACCAATGCTATGATGGTGGTCTGCTGATGAGAGCTGGACAAAGAGCTTGTATCACT 540
 Db 1894 GACCAATGCTATGATGGTGGTCTGCTGATGAGAGCTGGACAAAGAGCTTGTATCACT 1835
 QY 541 CAACAATCTGGAGTTTTCTGAGCGCGGATTTGGAGGGGATCGATGAGATTTCCCAAGC 600
 Db 1834 CAACAATCTGGAGTTTTCTGAGCGCGGATTTGGAGGGGATCGATGAGATTTCCCAAGC 1775
 QY 601 CGGCATCAACATTTGGCGAAGGCGCACCGATTCCAAAACCGCGAATACTAACCCATCAA 660
 Db 1774 CGGCATCAACATTTGGCGAAGGCGCACCGATTCCAAAACCGCGAATACTAACCCATCAA 1715
 QY 661 CATCATGTTGATGGCCATGCGGTGCTCATCAACTGCGCAGCAGCTGTTGATCCAGCGCA 720
 Db 1714 CATCATGTTGATGGCCATGCGGTGCTCATCAACTGCGCAGCAGCTGTTGATCCAGCGCA 1655
 QY 721 CACCTTGGGGCTGGACAGCGGGGTGACATGCTGCGCCGCAAAACCCACAGCGGAA 780
 Db 1654 CACCTTGGGGCTGGACAGCGGGGTGACATGCTGCGCCGCAAAACCCACAGCGGAA 1595

QY 781 CCAGATCAGGTTCCGCGAAGCGCGCAGCGGCGAAATCCACGTCGCGTGTCCGCTA 840
 Db 1594 CAGATCAGGTTCCGCGAAGCGCGCAGCGGCGAAATCCACGTCGCGTGTCCGCTA 1535
 QY 841 TTGCGCGCGAAGCGCGGATTAACAACAAACGCGTCCAAATACGATTCGGTTCACCA 900
 Db 1534 TTGCGCGCGAAGCGCGGATTAACAACAAACGCGTCCAAATACGATTCGGTTCACCA 1475
 QY 901 GGTACGACGATTCGATCAACATGGGCTTTACCAACCGCGTCTTACGACGCTCAC 960
 Db 1474 GGTACGACGATTCGATCAACATGGGCTTTACCAACCGCGTCTTACGACGCTCAC 1415
 QY 961 CTCACCGCGACCGGTTCCGCGTGTCAAGTGGCCACCGCGAAACCGCGGCGTGTG 1020
 Db 1414 CTCACCGCGACCGGTTCCGCGTGTCAAGTGGCCACCGCGAAACCGCGGCGTGTG 1355
 QY 1021 ATCGCGACGTTGGTTCGTTCTTCAATGATTCGTGGGCGTTCACCTTTTGTGAT 1080
 Db 1354 ATCGCGACGTTGGTTCGTTCTTCAATGATTCGTGGGCGTTCACCTTTTGTGAT 1295
 QY 1081 GCGGTCCTTTCGTCGATCAAGCGAAACCATACAGGTAAAGGATGACCGCGCGAT 1140
 Db 1294 GCGGTCCTTTCGTCGATCAAGCGAAACCATACAGGTAAAGGATGACCGCGCGAT 1235
 QY 1141 AATATCGACGATCGCGCGGCAATTGGACAAAGATCAACGCGCGAAGTGC CGGCGAT 1200
 Db 1234 AATATCGACGATCGCGCGGCAATTGGACAAAGATCAACGCGCGAAGTGC CGGCGAT 1175
 QY 1201 GAACAAAGACGTCAGAAATTAACACACGAGAGAACCGCATGAGTCTTCGCGGT 1260
 Db 1174 GAACAAAGACGTCAGAAATTAACACACGAGAGAACCGCATGAGTCTTCGCGGT 1115
 QY 1261 AATTCCTTGTATTAATCACCAGTACATTCGCGGTCCGATGACAGTAAGAACTGCGCCC 1320
 Db 1114 AATTCCTTGTATTAATCACCAGTACATTCGCGGTCCGATGACAGTAAGAACTGCGCCC 1055
 QY 1321 CAAAGACGATGATTAATGAAGATTCATGATCAACATGCTGATGACCTATGGAAGTACTTA 1380
 Db 1054 CAAAGACGATGATTAATGAAGATTCATGATCAACATGCTGATGACCTATGGAAGTACTTA 995
 QY 1381 AGTAAATGATGTTGTTTAAACATGTTTAAATATAGTTTCAATGAGTCAACCTCACTGGA 1440
 Db 994 AGTAAATGATGTTGTTTAAACATGTTTAAATATAGTTTCAATGAGTCAACCTCACTGGA 935
 QY 1441 CACTTTGCTCTCAATCATTTGATGAGGAGCTTCGAGGCGCTCTTACGCTTTCCTTCCAT 1500
 Db 934 CACTTTGCTCTCAATCATTTGATGAGGAGCTTCGAGGCGCTCTTACGCTTTCCTTCCAT 875
 QY 1501 TTCCCTCTCGCGGTGAGTCAGCGGTTAAAGCTCTCGAGCATCACTGGGTGAGTGT 1560
 Db 874 TTCCCTCTCGCGGTGAGTCAGCGGTTAAAGCTCTCGAGCATCACTGGGTGAGTGT 815
 QY 1561 GGTATCGCGACCCCAACCGGCAAGCAACGAGCGGTGAAGTCTTGTGCAAGCAGC 1620
 Db 814 GGTATCGCGACCCCAACCGGCAAGCAACGAGCGGTGAAGTCTTGTGCAAGCAGC 755
 QY 1621 GCGGAAATGTTGTTCTCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTCTGTA 1680
 Db 754 GCGGAAATGTTGTTCTCAAGCAGAACTAAAGCGCAACTATCTGAGCGCTTCTGTA 695
 QY 1681 AATCCGTTACCATGCCATCAAGCGAGATTCGATCCATGTTTCTCCCGTGT 1740
 Db 694 AATCCGTTACCATGCCATCAAGCGAGATTCGATCCATGTTTCTCCCGTGT 635
 QY 1741 CAACGAGTGTCTTTCGCGGTGAGCAACGCTCAGCTGCGCTTGAAGATGAAGCGCA 1800
 Db 634 CAACGAGTGTCTTTCGCGGTGAGCAACGCTCAGCTGCGCTTGAAGATGAAGCGCA 575
 QY 1801 CACATTATCTGTCGCGGTGAGATGTTTTAGAGCGGTAACCGCTGAAGCTAATCC 1860
 Db 574 CACATTATCTGTCGCGGTGAGATGTTTTAGAGCGGTAACCGCTGAAGCTAATCC 515
 QY 1861 CGTGGCGGATGAGTAGTAGAATTGGAAACATGCGCCACTGCGCATTTGCAACCC 1920

Db 514 CGTGGCGGATGTGAGTAGTAGAATTTGGAACCATGCGCCACTTGGCCATGCAACCC 455
 QY 1921 CTCATTGCGGATGCTACATGTTGATGGAACTAGATTTGGGCTGCGATGCCGCTTT 1980
 Db 454 CTCATTGCGGATGCTTACATGTTGATGGAACTAGATTTGGGCTGCGATGCCGCTTT 395
 QY 1981 AGCTTCGCTCCCAAGATGCTTCAAGACCGTGAACCGGCGCTGATGCTGCTCC 2040
 Db 394 AGCTTCGCTCCCAAGATGCTTCAAGACCGTGAACCGGCGCTGATGCTGCTGCTCC 335
 QY 2041 TTGCGGCGCAGGCGGCTATCCATTTGTCGCGGAAAGTTTTGGTGAAGCAATTCG 2100
 Db 334 TTGCGGCGCAGGCGGCTATCCATTTGTCGCGGAAAGTTTTGGTGAAGCAATTCG 275
 QY 2101 CCGAGGCTTGGTGGGAGCTTTCGGAACCGAGCTGCTCCATGCTAAAGCAGG 2160
 Db 274 CCGAGGCTTGGTGGGAGCTTTCGGAACCGAGCTGCTCCATGCTAAAGCAGG 215
 QY 2161 AGAAGTGTATCTCTCGATGAGATACCATTTGACACACCGATGATTTGGCAACGATGGCG 2220
 Db 214 AGAAGTGTATCTCTCGATGAGATACCATTTGACACACCGATGATTTGGCAACGATGGCG 155
 QY 2221 CTTGGAATCTAGATCTCTAGCTAGATCTCACAGAGCGGCTGTTGATGAGCAATTCAGGG 2280
 Db 154 CTTGGAATCTAGATCTCTAGCTAGATCTCACAGAGCGGCTGTTGATGAGCAATTCAGGG 95
 QY 2281 ATTGCGGCTTGTACTTCTGAAAGGTTTCAAGGTTTTCCTTCTTCCCGCGCAGGAA 2340
 Db 94 ATTGCGGCTTGTACTTCTGAAAGGTTTCAAGGTTTTCCTTCTTCCCGCGCAGGAA 35
 QY 2341 TTGGGCGCAGGAGTAGTAACACTTTCAGCAATGG 2374
 Db 34 TTGGGCGCAGGAGTAGTAACACTTTCAGCAATGG 1

RESULT 3
 AAH68528
 ID AAH68528 standard; DNA; 349980 BP.
 AC AAH68528;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 7063.
 XX
 KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KM organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EP1108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI; 2001-376931/40.
 XX
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Disclosure; SEQ ID NO 7063; 246bp + Sequence Listing; English.

xx The present invention provides a number of nucleotide and protein sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These are useful for identifying the mutation point of a gene derived from a mutant of corynebacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from corynebacterium, and identifying a homologue of a gene derived from corynebacterium. Corynebacterium bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a nucleic acid described in the exemplification of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office

xx
SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 U; 0 Other;

Query Match 100.0%; Score 2374; DB 5; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGATACCTCTTGGAGAAACCATGTACGATTCAGCTGCTGACATTTGTTGGCTCTGGAAAGGC 60
Db 127595 AGATACCTCTTGGAGAAACCATGTACGATTCAGCTGCTGACATTTGTTGGCTCTGGAAAGGC 127654

QY 61 TCTTTACGTGGGTATTTCTTCTTACGCTCCAGAGCTCACAGCGAGGCGGCTGAGTTTCAT 120
Db 127655 TCTTTACGTGGGTATTTCTTCTTACGCTCCAGAGCTCACAGCGAGGCGGCTGAGTTTCAT 127714

QY 121 GGCGGAGGAGGCTGCGGCTTCTGATTCATCAGCGAAGCTATTCATATTAATTCGTTG 180
Db 127715 GGCGGAGGAGGCTGCGGCTTCTGATTCATCAGCGAAGCTATTCATATTAATTCGTTG 127774

QY 181 GGTGAGAGAAACCGGCGATGACGCTGAGAACTTTGTCAGTCAAGCTGCCCAACATGTCT 240
Db 127775 GGTGAGAGAAACCGGCGATGACGCTGAGAACTTTGTCAGTCAAGCTGCCCAACATGTCT 127834

QY 241 TGGCGTCAATGCTTTTCCACCACTTGGCAGGCGCTGCTCACGGAATAATCTCGATGG 300
Db 127835 TGGCGTCAATGCTTTTCCACCACTTGGCAGGCGCTGCTCACGGAATAATCTCGATGG 127894

QY 301 AATTCCAGAGGTTCCCGGCGCAGCAGGCTAAGTCCCTGCTGAGGGCAGTGGTGAAGCT 360
Db 127895 AATTCCAGAGGTTCCCGGCGCAGCAGGCTAAGTCCCTGCTGAGGGCAGTGGTGAAGCT 127954

QY 361 GAACAATATTGATGCTCCGCAAGCTCAATGACATCCCGCAGAAACCGGCGAGTCACT 420
Db 127955 GAACAATATTGATGCTCCGCAAGCTCAATGACATCCCGCAGAAACCGGCGAGTCACT 128014

QY 421 TGGCGAGATGGGCTTGATGGGTGCTGCGGAGCAAGAGAGTACGCGCGGATACCGT 480
Db 128015 TGGCGAGATGGGCTTGATGGGTGCTGCGGAGCAAGAGAGTACGCGCGGATACCGT 128074

QY 481 GACCAATGATGATTTGTTGCTTCTGCTGCTGAGCAGCTGACACAGCCTTGATTCAC 540
Db 128075 GACCAATGATGATTTGTTGCTTCTGCTGCTGAGCAGCTGACACAGCCTTGATTCAC 128134

QY 541 CAACAACCTTGGAGTTTCTGACGCGGAGTTGGAGCGCATCGATGAGATTTCCACGAGC 600
Db 128135 CAACAACCTTGGAGTTTCTGACGCGGAGTTGGAGCGCATCGATGAGATTTCCACGAGC 128194

QY 601 CGGCATCAACATTTGGGCGAAGCCACCGATTTCCAAAACCGCGGAAAACCTAACCCATCAA 660
Db 128195 CGGCATCAACATTTGGGCGAAGCCACCGATTTCCAAAACCGCGGAAAACCTAACCCATCAA 128254

QY 661 CATCAGTTTCATGCGCATGGGTCACTCACTCCACGACGAGCTTGATCCAGCGCCA 720
Db 128255 CATCAGTTTCATGCGCATGGGTCACTCACTCCACGACGAGCTTGATCCAGCGCCA 128314

QY 721 CACCTTGGGCTGGACAGCGGCGGTGACAAATGCTGCGCGGAAAACCCACCGCGGAAA 780
Db 128315 CACCTTGGGCTGGACAGCGGCGGTGACAAATGCTGCGCGGAAAACCCACCGCGGAAA 128374

QY 781 CCAGATCAGGCTTGGCGGAACCGGCGCAGCGGCGGAAAATCCACCGTCCGGTGTGCGCGGTA 840

Db 128375 CAGATCAGGCTTGGCGGAAACCGGCGAGGCGGAAAATCCACGTCGCTGTCGCGGTA 128434

QY 841 TTGCGCGCGAGCGCGCGCGATAAACAAACACGCGTCCAAATACGATTCGGGTTCACCA 900
Db 128435 TTGCGCGCGAGCGCGCGCGATAAACAAACACGCGTCCAAATACGATTCGGGTTCACCA 128494

QY 901 GGTGAGCAGATTTGCCATCAAGTGGCTTACCCAAACCGGCTTATTCGACGCTCAC 960
Db 128495 GGTGAGCAGATTTGCCATCAAGTGGCTTATCCAAACCGGCTTATTCGACGCTCAC 128554

QY 961 CTCACCGCGACCGGCTTGGGCTGTCAGTGGCGACCGCGGAAACCGGCGGAGGCTGTC 1020
Db 128555 CTCACCGCGACCGGCTTGGGCTGTCAGTGGCGACCGCGGAAACCGGCGGAGGCTGTC 128614

QY 1021 ATCGGCGACGTTGGTTCGTTCTTCAATGATCTGTGGCGCTTCCACCTTGTGTGTCAT 1080
Db 128615 ATCGGCGACGTTGGTTCGTTCTTCAATGATCTGTGGCGCTTCCACCTTGTGTGTCAT 128674

QY 1081 GCGGCTTTTGGCTGCCATGAGCGGAAACCAATACAGTTAAGGATGCCACCCAGCGCAT 1140
Db 128675 GCGGCTTTTGGCTGCCATGAGCGGAAACCAATACAGTTAAGGATGCCACCCAGCGCAT 128734

QY 1141 AATATCGAGCAGCATCGGCGCGCATTTGGAACAAAGATCAACGCGGCAAGTGCAGCGCAT 1200
Db 128735 AATATCGAGCAGCATCGGCGCGCATTTGGAACAAAGATCAACGCGGCAAGTGCAGCGCAT 128794

QY 1201 GAACAAAAGAGCTGAGAAATTAACACACGAGAGAACCGGAATGAGTCTTTCGGGCTT 1260
Db 128795 GAACAAAAGAGCTGAGAAATTAACACACGAGAGAACCGGAATGAGTCTTTCGGGCTT 128854

QY 1261 AATCTCTTGTAAATCACCAGTACATCTCGGCTCGGATGACAGTAAAGAGTCTGGCCCC 1320
Db 128855 AATCTCTTGTAAATCACCAGTACATCTCGGCTCGGATGACAGTAAAGAGTCTGGCCCC 128914

QY 1321 CAAAAGCAGACCTGTAAATGAAGATTTCCATGATCAACCTCGTGAACCTTATGGAAGTACTTA 1380
Db 128915 CAAAAGCAGACCTGTAAATGAAGATTTCCATGATCAACCTCGTGAACCTTATGGAAGTACTTA 128974

QY 1381 AGTAAATGATGGTCTTAAATGATGTTTAAATAGTCTTATGATGACCCCATTCATCGGA 1440
Db 128975 AGTAAATGATGGTCTTAAATGATGTTTAAATAGTCTTATGATGACCCCATTCATCGGA 129034

QY 1441 CACTTTGCTCTCAATCATTGATGAGGAGCTTCCGAGGCGCTCTTAGCCCTTTCAT 1500
Db 129035 CACTTTGCTCTCAATCATTGATGAGGAGCTTCCGAGGCGCTCTTAGCCCTTTCAT 129094

QY 1501 TTCCCTCTTGGCGGTGAGTCAAGCGGTTAAAGCTCTCGAGCATCACTGTTGGTTCGAGTGT 1560
Db 129095 TTCCCTCTTGGCGGTGAGTCAAGCGGTTAAAGCTCTCGAGCATCACTGTTGGTTCGAGTGT 129154

QY 1561 GGTATCGCGCACCAACCGGCGCAAGCAACCGAGGCGGTGAGTCTTGTGCAAGCAGC 1620
Db 129155 GGTATCGCGCACCAACCGGCGCAAGCAACCGAGGCGGTGAGTCTTGTGCAAGCAGC 129214

QY 1621 GCGGAAAATGCTGCTGCAAGCAGAAACTAAAGCGCAACTATCTGAGCGCTTCTCTGA 1680
Db 129215 GCGGAAAATGCTGCTGCAAGCAGAAACTAAAGCGCAACTATCTGAGCGCTTCTCTGA 129274

QY 1681 AATCCGTTAACCATCGCCATCAACGAGATTCGCTATCCATGTTTCCTCCCGTGT 1740
Db 129275 AATCCGTTAACCATCGCCATCAACGAGATTCGCTATCCATGTTTCCTCCCGTGT 129334

QY 1741 CAAAGAGTACTCTTGTGGGTGAGCAACGCTCACTGCTGCGTTCGAGATGAAGCGCA 1800
Db 129335 CAAAGAGTACTCTTGTGGGTGAGCAACGCTCACTGCTGCGTTCGAGATGAAGCGCA 129394

QY 1801 CACATTATCTTCTGCGCGGTGAGATGTTTTTAGGAGCGGTAAACCGGTGAAGCTTAATCC 1860
Db 129395 CACATTATCTTCTGCGCGGTGAGATGTTTTTAGGAGCGGTAAACCGGTGAAGCTTAATCC 129454

QY 1861 CGTGGCGGATGTGAAGTAGTAGAACTTGGAAACCATGCGCACTTGGCCATTTGCAACCCC 1920

Db 129455 CGTGGCGGATGTGAAGTGTAGTAACTTGGAAACCATGCGCCACTTGGCCATGTCACACCC 129514
QY 1321 CTCATTGGGATGATCCCTACATGTTGATGGAACTAGATTGGGCTCGATGCCGTCTT 1380
Db 129515 CTCATTGGGATGATCCCTACATGTTGATGGAACTAGATTGGGCTCGATGCCGTCTT 129574
QY 1381 ACSCTTCGGTCCCAAGATGTGTTCAAGACCGTGCCTTGGACGGGGCGTCCATGGTCC 2040
Db 129575 ACSCTTCGGTCCCAAGATGTGTTCAAGACCGTGCCTTGGACGGGGCGTCCATGGTCC 129634
QY 2041 TGTGGGCGCGAGCGGCTATCCATTGTCCTCGTGGCGGAAGTCTTGGTGGAGCAATTCG 2100
Db 129635 TGTGGGCGCGAGCGGCTATCCATTGTCCTCGTGGCGGAAGTCTTGGTGGAGCAATTCG 129694
QY 2101 CCGAGGCTTGGTGGGGCTTCTCCGAAACCCCAAGCTGCTCCCATGCTAAAGCAGG 2160
Db 129695 CCGAGGCTTGGTGGGGCTTCTCCGAAACCCCAAGCTGCTCCCATGCTAAAGCAGG 129754
QY 2161 AGAAGTGTATCTCTCGATGAGATACCCATTGACACCGATGATTTGGCAACGATGGCG 2220
Db 129755 AGAAGTGTATCTCTCGATGAGATACCCATTGACACCGATGATTTGGCAACGATGGCG 129814
QY 2221 CTTGGAATCTAGATCTCTAGCTAGACTCAGAGCGGCTGTTGATGAGCAATCGAGG 2280
Db 129815 CTTGGAATCTAGATCTCTAGCTAGACTCAGAGCGGCTGTTGATGAGCAATCGAGG 129874
QY 2281 ATTGGCGCTTAGTTACTTCTGAAAAGTTTCAGGGTTTTCACCTTCTTCGCCCGCAGGAA 2340
Db 129875 ATTGGCGCTTAGTTACTTCTGAAAAGTTTCAGGGTTTTCACCTTCTTCGCCCGCAGGAA 129934
QY 2341 TTGGGCGAGGAGAGTAAACCTTCAGCAATGG 2374
Db 129935 TTGGGCGAGGAGAGTAAACCTTCAGCAATGG 129968

RESULT 4

AAAF71777
ID AAFA71777 standard; DNA; 993 BP.
XX
AC AAFA71777;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:49.
XX
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB000923.
XX
PR 25-JUN-1999; 99US-0141031P.
PR 01-JUL-1999; 99DE-01030476.
PR 02-JUL-1999; 99US-0142101P.
PR 08-JUL-1999; 99DE-01031415.
PR 08-JUL-1999; 99DE-01031418.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031435.
PR 08-JUL-1999; 99DE-01031443.
PR 08-JUL-1999; 99DE-01031453.
PR 08-JUL-1999; 99DE-01031457.

PR 08-JUL-1999; 99DE-01031465.
PR 08-JUL-1999; 99DE-01031478.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031541.
PR 08-JUL-1999; 99DE-01031573.
PR 08-JUL-1999; 99DE-01031592.
PR 08-JUL-1999; 99DE-01031632.
PR 08-JUL-1999; 99DE-01031634.
PR 08-JUL-1999; 99DE-01031636.
PR 08-JUL-1999; 99DE-01032125.
PR 08-JUL-1999; 99DE-01032126.
PR 08-JUL-1999; 99DE-01032130.
PR 08-JUL-1999; 99DE-01032186.
PR 08-JUL-1999; 99DE-01032206.
PR 08-JUL-1999; 99DE-01032227.
PR 08-JUL-1999; 99DE-01032228.
PR 08-JUL-1999; 99DE-01032229.
PR 08-JUL-1999; 99DE-01032230.
PR 08-JUL-1999; 99DE-01032232.
PR 14-JUL-1999; 99DE-01032926.
PR 14-JUL-1999; 99DE-01032928.
PR 14-JUL-1999; 99DE-01033004.
PR 14-JUL-1999; 99DE-01033005.
PR 14-JUL-1999; 99DE-01033006.
PR 12-AUG-1999; 99US-0148613P.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99DE-01040765.
PR 27-AUG-1999; 99DE-01040766.
PR 27-AUG-1999; 99DE-01040832.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041379.
PR 31-AUG-1999; 99DE-01041380.
PR 31-AUG-1999; 99DE-01041394.
PR 31-AUG-1999; 99DE-01041396.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042077.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042124.
PR 03-SEP-1999; 99DE-01042129.
PR 09-MAR-2000; 2000US-0187970P.
XX
XX (BADI) BASF AG.
XX PA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX PI WPI; 2001-137957/14.
XX XX P-PSDB; AAB79658.
XX DR Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway
XX PT proteins, useful for producing fine chemicals in microorganisms,
XX FT including organic acids, nonproteinogenic amino acids, and purine and
XX FT pyrimidine bases.
XX PS Claim 3; Page 226-228; 1737pp; English.
XX CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
XX CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
XX CC nucleic acids are useful for the production of fine chemicals in
XX CC microorganisms, including organic acids, nonproteinogenic amino acids,
XX CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
XX CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
XX CC vitamins, cofactors, polyketides and enzymes
XX
XX SQ Sequence 993 BP; 233 A; 256 C; 268 G; 236 T; 0 U; 0 Other;
Query Match 41.8%; Score 993; DB 4; Length 993;
Best Local Similarity 100.0%; Pred. No. 4.9e-299;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1321 CAAAGCAGACCTGTAATGAGATTTTCCATGATCACCATCGTGACCTATGGAAGTACTTA 1380
 Db 1 CAAAGCAGACCTGTAATGAGATTTTCCATGATCACCATCGTGACCTATGGAAGTACTTA 60
 QY 1381 AGTAAATGATTTGTTCTTAACATGTTTAAATATAGTTTCAAGACCCCATTTCACTGGA 1440
 Db 61 AGTAAATGATTTGTTCTTAACATGTTTAAATATAGTTTCAAGACCCCATTTCACTGGA 120
 QY 1441 CACTTTGCTCTCAATCATTCATGATGAGGAGCTTGAAGGCGCTCTTGAAGCCCTTTCCAT 1500
 Db 121 CACTTTGCTCTCAATCATTCATGATGAGGAGCTTGAAGGCGCTCTTGAAGCCCTTTCCAT 180
 QY 1501 TTCCCTCTCGGCGGTGAGTCAAGCTTTAAAGCTCTCGAGCATCACTGAGGCTGAGTGT 1560
 Db 181 TTCCCTCTCGGCGGTGAGTCAAGCTTTAAAGCTCTCGAGCATCACTGAGGCTGAGTGT 240
 QY 1561 GSTATCGCCAGCCCAACCGGCGCAAGACCAACGAGCGGCTGAGTCTTGTGCAAGCAGC 1620
 Db 241 GSTATCGCCAGCCCAACCGGCGCAAGACCAACGAGCGGCTGAGTCTTGTGCAAGCAGC 300
 QY 1621 GCGGAAATGCTGTTGCTGCAAGCAGAACTAAAGCGCACTATCTGAGCGCTTGTGCA 1680
 Db 301 GCGGAAATGCTGTTGCTGCAAGCAGAACTAAAGCGCACTATCTGAGCGCTTGTGCA 360
 QY 1681 AATCCCGTTAACAATGCGCATCAACGAGATTCGATATCAATGCTTTCCTCCGTTGT 1740
 Db 361 AATCCCGTTAACAATGCGCATCAACGAGATTCGATATCAATGCTTTCCTCCGTTGT 420
 QY 1741 CAACGAGGTAGCTTCTGGGCTGAGCAGCAGCTCAGCTCGCTTGGAGATGAGGCA 1800
 Db 421 CAACGAGGTAGCTTCTGGGCTGAGCAGCAGCTCAGCTCGCTTGGAGATGAGGCA 480
 QY 1801 CACATATCTCTGCTGCGGCTGAGATGTTTGAAGCGGTAACCCGCTGAAGCTAATCC 1860
 Db 481 CACATATCTCTGCTGCGGCTGAGATGTTTGAAGCGGTAACCCGCTGAAGCTAATCC 540
 QY 1861 GGTGCGGATGAGTGAAGTAGATAGAACTGGAACCATCGGCACTTGGCCTATGGAACCC 1920
 Db 541 GGTGCGGATGAGTGAAGTAGATAGAACTGGAACCATCGGCACTTGGCCTATGGAACCC 600
 QY 1921 CTCAATGCGGATGCTTACATGTTGATGGAATAGATGAGTGGCTGCGATGCGCTCTT 1980
 Db 601 CTCAATGCGGATGCTTACATGTTGATGGAATAGATGAGTGGCTGCGATGCGCTCTT 660
 QY 1981 ACCTTCGCTCCCAAGATGCTTCAAGACCGGTGACCTGGAACCGGCGCTGATGCTCC 2040
 Db 661 ACCTTCGCTCCCAAGATGCTTCAAGACCGGTGACCTGGAACCGGCGCTGATGCTCC 720
 QY 2041 TGTGGGCGCGAGCGCTATCCATGTTGCTGGCGGAAGGTTTGGTGAAGCAATTCG 2100
 Db 721 TGTGGGCGCGAGCGCTATCCATGTTGCTGGCGGAAGGTTTGGTGAAGCAATTCG 780
 QY 2101 CCGAGCGCTTGGTGGGAGCTTCTTCCGAAACCAAGCTGCTCCCATGCTAAAGCAGG 2160
 Db 781 CCGAGCGCTTGGTGGGAGCTTCTTCCGAAACCAAGCTGCTCCCATGCTAAAGCAGG 840
 QY 2161 AGAAGTATGCTTCCTGATGAGTACCAATGCAACCGATGATATGGCAACGATGGCG 2220
 Db 841 AGAAGTATGCTTCCTGATGAGTACCAATGCAACCGATGATATGGCAACGATGGCG 900
 QY 2221 CCTGGATAGTATCTCTAGCTAGATCAAGAGCGGCTGTTGATGAGCAATTCAGGG 2280
 Db 901 CCTGGATAGTATCTCTAGCTAGATCAAGAGCGGCTGTTGATGAGCAATTCAGGG 960
 QY 2281 ATTGGCGCTTATGTTACTTCTGAAAGGTTTCAG 2313
 Db 961 ATTGGCGCTTATGTTACTTCTGAAAGGTTTCAG 993

RESULT 5
 AAS96096
 ID AAS96096 standard; DNA; 993 BP.
 XX

AAS96096;
 26-FEB-2002 (first entry)
 C. glutamicum gene #21 encoding metabolic pathway protein.
 Metabolic pathway protein; MP; lysine biosynthesis pathway;
 methionine biosynthesis pathway; large-scale production of fine chemical;
 Corynebacterium diphtheriae; diphtheria; ds.
 Corynebacterium glutamicum.
 WO200166573-A2.
 13-SEP-2001.
 22-DEC-2000; 2000WO-IB002035.
 09-MAR-2000; 2000US-0187970P.
 23-JUN-2000; 2000US-00606740.
 (BADI) BASF AG.
 Pompejua M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 Kim J, Lee H, Hwang B;
 WPI: 2001-582269/65.
 P-PSDB; AAU71886.
 Nucleic acids encoding metabolic pathway proteins from Corynebacterium
 glutamicum, useful for producing methionine and lysine in Corynebacterium
 and Brevibacterium.
 Disclosure; Page 207-208; 316pp; English.
 The present invention relates to the isolation of novel Corynebacterium
 glutamicum genes encoding metabolic pathways (MP) proteins (AAU71863-
 AAU71922). The metabolic pathway proteins of the invention include
 enzymes involved in the lysine and methionine biosynthetic pathways. The
 polynucleotide sequences of the invention can be used for the large-scale
 production and/or modulation of expression of fine chemicals such as
 lysine and methionine. The sequences of the invention may be used to
 identify C. glutamicum and related organisms e.g. C. diphtheriae in a
 subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum
 genes encoding the novel metabolic pathway proteins of the invention
 Sequence 993 BP; 233 A; 256 C; 268 G; 236 T; 0 U; 0 Other;
 Query Match 41.8%; Score 993; DB 4; Length 993;
 Best Local Similarity 100.0%; Pred. No. 4.9e-299;
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1321 CAAAGCAGACCTGTAATGAGATTTTCCATGATCACCATCGTGACCTATGGAAGTACTTA 1380
 Db 1 CAAAGCAGACCTGTAATGAGATTTTCCATGATCACCATCGTGACCTATGGAAGTACTTA 60
 QY 1381 AGTAAATGATTTGTTCTTAACATGTTTAAATATAGTTTCAAGACCCCATTTCACTGGA 1440
 Db 61 AGTAAATGATTTGTTCTTAACATGTTTAAATATAGTTTCAAGACCCCATTTCACTGGA 120
 QY 1441 CACTTTGCTCTCAATCATTCATGATGAGGAGCTTGAAGGCGCTCTTGAAGCCCTTTCCAT 1500
 Db 121 CACTTTGCTCTCAATCATTCATGATGAGGAGCTTGAAGGCGCTCTTGAAGCCCTTTCCAT 180
 QY 1501 TTCCCTCTCGGCGGTGAGTCAAGCTTTAAAGCTCTCGAGCATCACTGAGGCTGAGTGT 1560
 Db 181 TTCCCTCTCGGCGGTGAGTCAAGCTTTAAAGCTCTCGAGCATCACTGAGGCTGAGTGT 240
 QY 1561 GSTATCGCCAGCCCAACCGGCGCAAGACCAACGAGCGGCTGAGTCTTGTGCAAGCAGC 1620
 Db 241 GSTATCGCCAGCCCAACCGGCGCAAGACCAACGAGCGGCTGAGTCTTGTGCAAGCAGC 300
 QY 1621 GCGGAAATGCTGTTGCTGCAAGCAGAACTAAAGCGCACTATCTGAGCGCTTGTGCA 1680

Db 301 GCGGAAATGTTGTCGACGCGAATCTTAAGCGCACTATCTGAGCGCTTGTCTGA 360
QY 1681 AATCCGCTTAACCATCCCATCAAGCGAGATTCGCTATCCACATGGTTTCTCCGCTGT 1740
Db 361 AATCCGCTTAACCATCCCATCAAGCGAGATTCGCTATCCACATGGTTTCTCCGCTGT 420
QY 1741 CAACGAGGTAGCTTCTTGGGCTGGAGCAACGCTCAAGCTGCGCTTGGAGATCAAGCGCA 1800
Db 421 CAACGAGGTAGCTTCTTGGGCTGGAGCAACGCTCAAGCTGCGCTTGGAGATCAAGCGCA 480
QY 1801 CACATTAATCTTCTGCTGGGCTGGAGATGTTTGGAGCGGTAAACCGCTGAAGCTTAATCC 1860
Db 481 CACATTAATCTTCTGCTGGGCTGGAGATGTTTGGAGCGGTAAACCGCTGAAGCTTAATCC 540
QY 1861 CGTGGCGGATGTAAGTAGTAGAATCTTGAACCATCGCGCACTTGGCCATTCGAACCC 1920
Db 541 CGTGGCGGATGTAAGTAGTAGAATCTTGAACCATCGCGCACTTGGCCATTCGAACCC 600
QY 1921 CTCATTTGCGGATGCTACATGTTGATGGCAAACTAGATTGGGCTGCGATGCCGCTTT 1980
Db 601 CTCATTTGCGGATGCTACATGTTGATGGCAAACTAGATTGGGCTGCGATGCCGCTTT 660
QY 1981 ACGCTTCGCTCCCAAGATGCTTCAAGACCGTGAACCTGCGACGGGCGCTGCGATGCTCC 2040
Db 661 ACGCTTCGCTCCCAAGATGCTTCAAGACCGTGAACCTGCGACGGGCGCTGCGATGCTCC 720
QY 2041 TGTGGGCGCGAGCGCGTATCCATGTTGCTCCGTCGCGGAGGTTTGTGGAGCAATTCG 2100
Db 721 TGTGGGCGCGAGCGCGTATCCATGTTGCTCCGTCGCGGAGGTTTGTGGAGCAATTCG 780
QY 2101 CCGAGGCTTGTGTTGGGCACTTCTTCCGCAACCCAAAGCTGCTCCCATGCTTAAAGCAGG 2160
Db 781 CCGAGGCTTGTGTTGGGCACTTCTTCCGCAACCCAAAGCTGCTCCCATGCTTAAAGCAGG 840
QY 2161 AGAAGTATCTCTCGATGAGATCCCATTTGACACACCGATGATTGGCAAGGATGGCG 2220
Db 841 AGAAGTATCTCTCGATGAGATCCCATTTGACACACCGATGATTGGCAAGGATGGCG 900
QY 2221 CCGTGAATCTAGATCTCTAGCTAGATCAACAGCGCGCTGTTGATGCGCAATTCGAGG 2280
Db 901 CCGTGAATCTAGATCTCTAGCTAGATCAACAGCGCGCTGTTGATGCGCAATTCGAGG 2340
QY 2281 ATTGCGGCTTAGTACTTCTGAAAGGTTTCAAG 2313
Db 961 ATTGCGGCTTAGTACTTCTGAAAGGTTTCAAG 993

RESULT 6

AAH68421
ID AAH68421 standard; DNA; 870 BP.
XX AC AAH68421;
XX AC
DT 26-SEP-2001 (first entry)
XX C glutamicum coding sequence fragment SEQ ID NO: 3456.
DE DE
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; ds.
XX Corynebacterium glutamicum.
XX EP1108790-A2.
XX 20-JUN-2001.
XX 18-DEC-2000; 2000EP-00127688.
XX 16-DEC-1999; 99JP-00377484.
XX 07-APR-2000; 2000JP-00159162.
XX 03-AUG-2000; 2000JP-00280988.
XX

PA (KYOM) KYOMA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
DR P-PSDB; AAG93202.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX Claim 1; SEQ ID NO 3456; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 870 BP; 192 A; 236 C; 245 G; 197 T; 0 U; 0 Other;
Query Match 36.6%; Score 870; DB 5; Length 870;
Best Local Similarity 100.0%; Pred. No. 1.3e-260;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1421 ATGAACCCCATTAACATGACACTTTGCTCTCAATCATGATGAAGCAGCTTCGAAGC 1480
Db 1 ATGAACCCCATTAACATGACACTTTGCTCTCAATCATGATGAAGCAGCTTCGAAGC 60
QY 1481 GCCTCTCTTACGCTTTTCCATTTTCCCTCTCGGCTGAGTCAGCGGTTTAAAGCTCTCGAG 1540
Db 61 GCCTCTCTTACGCTTTTCCATTTTCCCTCTCGGCTGAGTCAGCGGTTTAAAGCTCTCGAG 120
QY 1541 CATCACTGCTGCTGAGTGTGTTGATTCGCGCAACCGGCTAAAGCAACCGAAGCGGT 1600
Db 121 CATCACTGCTGCTGAGTGTGTTGATTCGCGCAACCGGCTAAAGCAACCGAAGCGGT 180
QY 1601 GAGTCTCTTGTCAAGCAGCGCGGAAATGTTGTTGCTGCAAGCAGAACTTAAAGCGCAA 1660
Db 181 GAGTCTCTTGTCAAGCAGCGCGGAAATGTTGTTGCTGCAAGCAGAACTTAAAGCGCAA 240
QY 1661 CTATCTGAGCGCTTGTGAAATCCCGTTTAAACCATTCGCTATCAACGAGATTCGCTATCC 1720
Db 241 CTATCTGAGCGCTTGTGAAATCCCGTTTAAACCATTCGCTATCAACGAGATTCGCTATCC 300
QY 1721 ACATGTTTCTCCGCTTCAAGAGGTAGCTTCTTGGGCTGGAGCAACCGCTCAGCGTG 1780
Db 301 ACATGTTTCTCCGCTTCAAGAGGTAGCTTCTTGGGCTGGAGCAACCGCTCAGCGTG 360
QY 1781 CGCTTGAAGATGAAGCGCACATTTCTTCTGCTGCGGCTGGAGATGTTTGAAGCGG 1840
Db 361 CGCTTGAAGATGAAGCGCACATTTCTTCTGCTGCGGCTGGAGATGTTTGAAGCGG 420
QY 1841 GTAAACCGGTGAAGCTTAATCCCGTGGGCGGATGTAAGTAGTAGAACTTGAACCATCGCG 1900
Db 421 GTAAACCGGTGAAGCTTAATCCCGTGGGCGGATGTAAGTAGTAGAACTTGAACCATCGCG 480
QY 1901 CACTTGGCCATTTGAACCCCTCATTTGGGGATCGCTTACATGTTGATGGGAATAGAT 1960
Db 481 CACTTGGCCATTTGAACCCCTCATTTGGGGATCGCTTACATGTTGATGGGAATAGAT 540
QY 1961 TGGGCTGCGATCGCGCTTACCTTTCGCTCCCAAGATGTTGTTCAAGACCGTGACCTG 2020
Db 541 TGGGCTGCGATCGCGCTTACCTTTCGCTCCCAAGATGTTGTTCAAGACCGTGACCTG 600

QY 2021 GACGGCGCGTGCATGCTCTCTGGGGCGCAGCGCGTATCCATTTGTCGGCGGAA 2080
 Db 601 GACGGCGCGTGCATGCTCTCTGGGGCGCAGCGCGTATCCATTTGTCGGCGGAA 660
 QY 2081 GCTTTTGGTGAAGCAATTCGCGAGGCTTGGTGGGACCTTCTTCCGAAACCCAAAGCT 2140
 Db 661 GCTTTTGGTGAAGCAATTCGCGAGGCTTGGTGGGACCTTCTTCCGAAACCCAAAGCT 720
 QY 2141 GCTCCCATGCTAAACAGAGAGAGTGCATCTCTCTGATGATACCATTCGACACACCG 2200
 Db 721 GCTCCCATGCTAAACAGAGAGAGTGCATCTCTCTGATGATACCATTCGACACACCG 780
 QY 2201 ATGATTTGGCAACGATGGCGCTGGAATCTAGATCTCTAGCTAGACTCAGACGCGGTC 2260
 Db 781 ATGATTTGGCAACGATGGCGCTGGAATCTAGATCTCTAGCTAGACTCAGACGCGGTC 840
 QY 2261 GTTATGATCAGCAATCGAGGATTCGGGCT 2290
 Db 841 GTTATGATCAGCAATCGAGGATTCGGGCT 870

RESULT 7

AAE71779/C
 ID AAE71779 standard; DNA; 822 BP.

XX AC AAE71779;

XX AC AAE71779;

DT 30-APR-2001 (first entry)

XX DE Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:53.

XX KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
 KW fine chemical production; microorganism; organic acid; nucleoside;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme; ds.

XX OS Corynebacterium glutamicum.

XX PN WC200100843-A2.

XX PD 04-JAN-2001.

XX PF 23-JUN-2000; 2000WO-1B000923.

XX PR 25-JUN-1999; 99US-0141031P.

PR 01-JUL-1999; 99DE-01030476.

PR 02-JUL-1999; 99US-0142101P.

PR 08-JUL-1999; 99DE-01031415.

PR 08-JUL-1999; 99DE-01031418.

PR 08-JUL-1999; 99DE-01031419.

PR 08-JUL-1999; 99DE-01031420.

PR 08-JUL-1999; 99DE-01031424.

PR 08-JUL-1999; 99DE-01031428.

PR 08-JUL-1999; 99DE-01031434.

PR 08-JUL-1999; 99DE-01031435.

PR 08-JUL-1999; 99DE-01031443.

PR 08-JUL-1999; 99DE-01031453.

PR 08-JUL-1999; 99DE-01031457.

PR 08-JUL-1999; 99DE-01031455.

PR 08-JUL-1999; 99DE-01031478.

PR 08-JUL-1999; 99DE-01031510.

PR 08-JUL-1999; 99DE-01031541.

PR 08-JUL-1999; 99DE-01031573.

PR 08-JUL-1999; 99DE-01031592.

PR 08-JUL-1999; 99DE-01031632.

PR 08-JUL-1999; 99DE-01031634.

PR 08-JUL-1999; 99DE-01031636.

PR 09-JUL-1999; 99DE-01032125.

PR 09-JUL-1999; 99DE-01032126.

PR 09-JUL-1999; 99DE-01032130.

PR 09-JUL-1999; 99DE-01032186.

PR 09-JUL-1999; 99DE-01032206.

PR 09-JUL-1999; 99DE-01032227.
 PR 09-JUL-1999; 99DE-01032228.
 PR 09-JUL-1999; 99DE-01032229.
 PR 09-JUL-1999; 99DE-01032230.
 PR 14-JUL-1999; 99DE-01032922.
 PR 14-JUL-1999; 99DE-01032926.
 PR 14-JUL-1999; 99DE-01032928.
 PR 14-JUL-1999; 99DE-01033004.
 PR 14-JUL-1999; 99DE-01033005.
 PR 14-JUL-1999; 99DE-01033006.
 PR 12-AUG-1999; 99US-0148613P.
 PR 27-AUG-1999; 99DE-01040764.
 PR 27-AUG-1999; 99DE-01040765.
 PR 27-AUG-1999; 99DE-01040766.
 PR 27-AUG-1999; 99DE-01040832.
 PR 31-AUG-1999; 99DE-01041378.
 PR 31-AUG-1999; 99DE-01041379.
 PR 31-AUG-1999; 99DE-01041380.
 PR 31-AUG-1999; 99DE-01041394.
 PR 31-AUG-1999; 99DE-01041396.
 PR 03-SEP-1999; 99DE-01042076.
 PR 03-SEP-1999; 99DE-01042077.
 PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042086.
 PR 03-SEP-1999; 99DE-01042087.
 PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042124.
 PR 03-SEP-1999; 99DE-01042129.
 PR 09-MAR-2000; 2000US-0187970P.
 XX (BADI) BASF AG.
 XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 PI WPI; 2001-137957/14.
 XX P-PSDB; AAB79660.
 DR Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway
 PT proteins, useful for producing fine chemicals in microorganisms,
 PT including organic acids, nonproteinogenic amino acids, and purine and
 PT pyrimidine bases.
 XX Claim 3; Page 233-234; 1737pp; English.
 PS AAE71753 to AAE72330 encode the Corynebacterium glutamicum metabolic
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
 CC nucleic acids are useful for the production of fine chemicals in
 CC microorganisms, including organic acids, nonproteinogenic amino acids,
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
 CC vitamins, cofactors, polyketides and enzymes
 XX Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;
 SQ Query Match 34.6%; Score 822; DB 4; Length 822;
 Best Local Similarity 100.0%; Pred. No. 1.2e-245;
 Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 629 GATTCGAAACCGCGGAAACCTAAACCCATCAATCAGTTTGTATGCGCAATCGGTGATC 688
 Db 822 GATTCGAAACCGCGGAAACCTAAACCCATCAATCAGTTTGTATGCGCAATCGGTGATC 763
 QY 689 ACACTGCGACGACGAGTTGATCCAGCGCCACCTTTGGGGCTTGACAGCGGGGTGAC 748
 Db 762 ACACTGCGACGACGAGTTGATCCAGCGCCACCTTTGGGGCTTGACAGCGGGGTGAC 703
 QY 749 AATGCTGCTGCGCGGAAACCCACCGCGGAAACCAAGATCAGGCTTTCGCGAACCGCGCA 808
 Db 702 AATGCTGCTGCGCGGAAACCCACCGCGGAAACCAAGATCAGGCTTTCGCGAACCGCGCA 643
 QY 809 GCGCGGAAATTCACCGTCCGGTGTTCGCGGTATTCGCGCGGCGGCGGATTAACACA 868

642 GCGCGAATCCACCGTCGCGTGTGCGCGTATTGCGCGCGAGCGCGCGATAACACA 583
 869 AACCGTCCAAATACGATTCGGGTTCAACAGGTAGACAGATTCGATCAACATGGGC 928
 582 AACCGTCCAAATACGATTCGGGTTCAACAGGTAGACAGATTCGATCAACATGGGC 523
 929 TTTACCCAAACCCGCTGCTTATCGAGCGTCACTCTCAACCGCGACCGGTTGCGCGTCA 988
 522 TTTACCCAAACCCGCTGCTTATCGAGCGTCACTCTCAACCGCGACCGGTTGCGCGTCA 463
 989 GTGGCCACCGCGCAACCGCCAAAGCGGTGTCATCGGCAACCGGTTGCTTCTTCA 1048
 462 GTGGCCACCGCGCAACCGCCAAAGCGGTGTCATCGGCAACCGGTTGCTTCTTCA 403
 1049 ATGATCTGTGGCGTCCACCTGTTTGTGATGCGCTTTCGTCGCGCATGCGCAAC 1108
 402 ATGATCTGTGGCGTCCACCTGTTTGTGATGCGCTTTCGTCGCGCATGCGCAAC 343
 1109 CATACAGGTAAAGCGATGCGCACCCAGCGCATATATCGAGCAGCATGCGCGCGCATTTG 1168
 342 CATACAGGTAAAGCGATGCGCACCCAGCGCATATATCGAGCAGCATGCGCGCGCATTTG 283
 1169 GACAAAGATCAACCGCGATGCGCGCGATGAAACAAAGACCGTCAAGAAATTAACAC 1228
 282 GACAAAGATCAACCGCGATGCGCGCGATGAAACAAAGACCGTCAAGAAATTAACAC 223
 1229 ACAGAGAAACCGCAATGAGTCTTGTGATGCGCTTTCGTCGCGCATGCGCAAC 1288
 222 ACAGAGAAACCGCAATGAGTCTTGTGATGCGCTTTCGTCGCGCATGCGCAAC 163
 1289 TGGGTCGCGATGAGCAAGTAAAGACCTGGCGCCCAAGACAGACCTGTAATGAAGATTTC 1348
 162 TGGGTCGCGATGAGCAAGTAAAGACCTGGCGCCCAAGACAGACCTGTAATGAAGATTTC 103
 1349 ATGATCAGCATCGTGACCTATGAGAGTACTTAAGTAAATGATGGTCTTCAATGATT 1408
 102 ATGATCAGCATCGTGACCTATGAGAGTACTTAAGTAAATGATGGTCTTCAATGATT 43
 1409 TAATATAGCTTCATGACCCCACTTCACTGGACACTTTGCTC 1450
 42 TAATATAGCTTCATGACCCCACTTCACTGGACACTTTGCTC 1

RESULT 8
 AAS96098/c
 ID AAS96098 standard; DNA; 822 BP.
 XX AAS96098;
 AC AAS96098;
 XX 26-FEB-2002 (first entry)
 DT C. glutamicum gene #23 encoding metabolic pathway protein.
 DE Metabolic pathway protein; MP; lysine biosynthesis pathway;
 XX methionine biosynthesis pathway; large-scale production of fine chemical;
 KW Corynebacterium diphtheriae; diphtheria; ds.
 XX OS Corynebacterium glutamicum.
 XX WO200166573-A2.
 PN 13-SEP-2001.
 PD 22-DEC-2000; 2000WO-IB002035.
 XX 09-MAR-2000; 2000US-0187970P.
 PR 23-JUN-2000; 2000US-00606740.
 XX (BADI) BASF AG.
 XX Pompeius M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 PI Kim J, Lee H, Hwang B;
 XX

DR WPI: 2001-582269/65.
 DR P-PSDB; AAU71888.
 XX Nucleic acids encoding metabolic pathway proteins from Corynebacterium
 PT glutamicum, useful for producing methionine and lysine in Corynebacterium
 PT and Brevibacterium.
 XX Disclosure; Page 214-215; 316pp; English.
 XX The present invention relates to the isolation of novel Corynebacterium
 CC glutamicum genes encoding metabolic pathway (MP) proteins (AAU71863-
 CC AAU71922). The metabolic pathway proteins of the invention include
 CC enzymes involved in the lysine and methionine biosynthetic pathways. The
 CC polynucleotide sequences of the invention can be used for the large-scale
 CC production and/or modulation of expression of fine chemicals such as
 CC lysine and methionine. The sequences of the invention may be used to
 CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a
 CC subject to detect diphtheria. AAS96073-AAS96132 represent C. glutamicum
 CC genes encoding the novel metabolic pathway proteins of the invention
 XX
 SQ Sequence 822 BP; 167 A; 192 C; 246 G; 217 T; 0 U; 0 Other;
 Query Match 34.6%; Score 822; DB 4; Length 822;
 Best Local Similarity 100.0%; Pred. No. 1.2e-245;
 Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 629 GATTCCAAACCCGCGAAACTAACCATCAACATCAGTTTGTGATGGCCAAATGGCGTCATC 688
 DB 822 GATTCCAAACCCGCGAAACTAACCATCAACATCAGTTTGTGATGGCCAAATGGCGTCATC 763
 QY 689 ACAACTGCCACGACGAGTGTGATCCAGCGCCACACCTTTGGGGCTGAGCAGCGCGCTGAC 748
 DB 762 ACAACTGCCACGACGAGTGTGATCCAGCGCCACACCTTTGGGGCTGAGCAGCGCGCTGAC 703
 QY 749 AATGCTGTGCGCGAAGAACCCACAGCAGCGGAAACAGATCAGGTTGCGCGAAGCGGCCA 808
 DB 702 AATGCTGTGCGCGAAGAACCCACAGCAGCGGAAACAGATCAGGTTGCGCGAAGCGGCCA 643
 QY 809 GCGCGCAAAATCCACCGTCCGCGTGTGCGCGTATTGCGCGCGACCGCCGCGATAAACACA 868
 DB 642 GCGCGCAAAATCCACCGTCCGCGTGTGCGCGTATTGCGCGCGACCGCCGCGATAAACACA 583
 QY 869 AACCGTCCAAATACGATTCGGGTTCAACAGGTAGACAGATTCGATCAACATGGGC 928
 DB 582 AACCGTCCAAATACGATTCGGGTTCAACAGGTAGACAGATTCGATCAACATGGGC 523
 QY 929 TTTACCCAAACCCGCTGCTTATCGAGCGTCACTCTCAACCGCGACCGGTTGCGCGTCA 988
 DB 522 TTTACCCAAACCCGCTGCTTATCGAGCGTCACTCTCAACCGCGACCGGTTGCGCGTCA 463
 QY 989 GTGGCCACCGCGCAACCGCCAAAGCGGTGTCATCGGCAACCGGTTGCTTCTTCA 1048
 DB 462 GTGGCCACCGCGCAACCGCCAAAGCGGTGTCATCGGCAACCGGTTGCTTCTTCA 403
 QY 1049 ATGATCTGTGGCGTCCACCTGTTTGTGATGCGCTTTCGTCGCGCATGCGCAAC 1108
 DB 402 ATGATCTGTGGCGTCCACCTGTTTGTGATGCGCTTTCGTCGCGCATGCGCAAC 343
 QY 1109 CATACAGGTAAAGCGATGCGCACCCAGCGCATATATCGAGCAGCATGCGCGCGCATTTG 1168
 DB 342 CATACAGGTAAAGCGATGCGCACCCAGCGCATATATCGAGCAGCATGCGCGCGCATTTG 283
 QY 1169 GACAAAGATCAACCGCGATGCGCGCGATGAAACAAAGACCGTCAAGAAATTAACAC 1228
 DB 282 GACAAAGATCAACCGCGATGCGCGCGATGAAACAAAGACCGTCAAGAAATTAACAC 223
 QY 1229 ACAGAGAAACCGCAATGAGTCTTGTGATGCGCTTTCGTCGCGCATGCGCAAC 1288
 DB 222 ACAGAGAAACCGCAATGAGTCTTGTGATGCGCTTTCGTCGCGCATGCGCAAC 163
 QY 1289 TGGGTCGCGATGAGCAAGTAAAGACCTGGCGCCCAAGACAGACCTGTAATGAAGATTTC 1348
 DB 162 TGGGTCGCGATGAGCAAGTAAAGACCTGGCGCCCAAGACAGACCTGTAATGAAGATTTC 103

QY 1349 ATGATCACCACCTGACCTATGGAAGTACTTAAAGTAAATGATTGGTTCTTAACATGCTT 1408
DB 102 ATGATCACCACCTGACCTATGGAAGTACTTAAAGTAAATGATTGGTTCTTAACATGCTT 43
QY 1409 TAAATAGCTTCATGAACCCCAATCACTGGACACTTTGCTC 1450
DB 42 TAAATAGCTTCATGAACCCCAATCACTGGACACTTTGCTC 1

RESULT 9

ACC80941/c
ID ACC80941 standard; DNA; 711 BP.

XX AC ACC80941;
XX 27-OCT-2003 (revised)
DT 11-AUG-2003 (first entry)
XX
DE LysE protein encoding sequence.
KW L-lysine; L-arginine; LysE; ds.
OS Corynebacterium glutamicum.
XX
FH Key Location/Qualifiers
FT CDS 1..711
FT /*tag= a
FT /product= "lysE protein"
XX

FN EPI266966-A2.

XX 18-DEC-2002.

PF 05-JUN-2002; 2002EP-00012539.

PR 12-JUN-2001; 2001JP-00177075.

XX (AJIN) AJINOMOTO CO INC.

PI Gunji Y, Yasueda H;

XX WPI; 2003-241171/24.

DR P-PSDB; ABR58213.

XX
PT Novel DNA encoding variant of LysE protein from a coryneform bacterium,
PT when introduced into methanol assimilating bacterium, facilitates
PT excretion of L-lysine and/or L-arginine to outside of a cell.

XX Example 1; Page 17-18; 23pp; English.

XX The present invention relates to DNA encoding variants of protein with
CC loop region and six hydrophobic helices which facilitates excretion of L-
CC lysine and/or L-arginine to outside of cell of a methanol assimilating
CC bacterium when introduced into the bacterium. The method is used for
CC encoding a protein which facilitates excretion of L-lysine, L-arginine or
CC both of these L-amino acids to outside of a cell of a methanol
CC assimilating bacterium when DNA of the method is introduced into the
CC bacterium. The present sequence represents a lysE protein from
CC Brevibacterium lactofermentum encoding sequence. (Updated on 27-OCT-2003
CC to standardise OS field)

XX Sequence 711 BP; 135 A; 173 C; 222 G; 181 T; 0 U; 0 Other;

SQ Query Match 29.9%; Score 711; DB 8; Length 711;

Best Local Similarity 100.0%; Pred. No. 5,7e-211;

Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 CTAACCCATCAACATGAGTTGATGGCCATGCGGTGATCACTGCGACGACGTT 708

DB 711 CTAACCCATCAACATGAGTTGATGGCCATGCGGTGATCACTGCGACGACGTT 652

QY 709 GATCCAGCGCCACACCTTGGGGCTGGACAGCGGGGTGACATGCTGCGCGCGAAACC 768

DB 651 GATCCAGCGCCACACCTTGGGGCTGGACAGCGGGGTGACATGCTGCTGGCGCGAAACC 592
QY 769 CACACAGCGGGAACACAGATCAGGCTTGGCGGAACCGCGCAGCGGGAATAATCCACCGTCC 828
DB 591 CACACAGCGGGAACACAGATCAGGCTTGGCGGAACCGCGCAGCGGGAATAATCCACCGTCC 532
QY 829 GGTGTGCGCGTATTGGCGCGCGACCGCGGATATAACACAAACCGCTCCAAATACGCAAT 888
DB 531 GGTGTGCGCGTATTGGCGCGCGACCGCGGATATAACACAAACCGCTCCAAATACGCAAT 472
QY 889 CGGGTTCAACACAGGTGACGACGATTCATCAATGAGGCTTTACCCAAAACCGCGTCTT 948
DB 471 CGGGTTCAACACAGGTGACGACGATTCATCAATGAGGCTTTACCCAAAACCGCGTCTT 412
QY 949 ATCGACGCTCACCCTCCACCGGCAACCGGTTGGCGGTGTCAGTGGGCCACCGCGCAACCGCC 1008
DB 411 ATCGACGCTCACCCTCCACCGGCAACCGGTTGGCGGTGTCAGTGGGCCACCGCGCAACCGCC 352
QY 1009 CAAAGCGGTGTATCGGGGACCGGTTGGTTCTTCTTCAATGATCTGTGGCGCTTCCAC 1068
DB 351 CAAAGCGGTGTATCGGGGACCGGTTGGTTCTTCTTCAATGATCTGTGGCGCTTCCAC 292
QY 1069 CTGTGTTGTGATGCGGTCTTTCGTCGCATGACGCGCAACCAATACAGGTAAGCGATGCC 1128
DB 291 CTGTGTTGTGATGCGGTCTTTCGTCGCATGACGCGCAACCAATACAGGTAAGCGATGCC 232
QY 1129 ACCCAGCGCATATATCGGACGACGATCGGCGGCGCATTTGGACAAAGATCAACGCCAA 1188
DB 231 ACCCAGCGCATATATCGGACGACGATCGGCGGCGCATTTGGACAAAGATCAACGCCAA 172
QY 1189 GGTGCGCGCATGACAAACAAAGACGTGAGAAATTAACACAGAGAAACCGCAATGAG 1248
DB 171 GGTGCGCGCATGACAAACAAAGACGTGAGAAATTAACACAGAGAAACCGCAATGAG 112
QY 1249 TCCTTCGCGCTTAATTCCTTGTGTTTAATCAACAGTACATTTCTGCGGTCCGATGACAGTAA 1308
DB 111 TCCTTCGCGCTTAATTCCTTGTGTTTAATCAACAGTACATTTCTGCGGTCCGATGACAGTAA 52
QY 1309 AAGACTGCGCCCGCCAAAGCAGACCTGTAAATGAAGATTTCCATGATCACCAT 1359
DB 51 AAGACTGCGCCCGCCAAAGCAGACCTGTAAATGAAGATTTCCATGATCACCAT 1

RESULT 10

AAH68420/c

ID AAH68420 standard; DNA; 708 BP.

XX AC AAH68420;

XX DT 26-SEP-2001 (first entry)

XX C glutamicum coding sequence fragment SEQ ID NO: 3455.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; ds.

XX Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.
DR P-PSDB; AAG93201.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX Claim 1; SEQ ID NO 3455; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutanicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analyzing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 708 BP; 134 A; 173 C; 221 G; 180 T; 0 U; 0 Other;
Query Match 29.8%; Score 708; DB 5; Length 708;
Best Local Similarity 100.0%; Pred. No. 4.9e-210;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 652 ACCGATCAATCATGTTGATGCGGTCATCATCACTGCCAGCAGCGTTGAT 711
DB 708 ACCGATCAATCATGTTGATGCGGTCATCATCACTGCCAGCAGCGTTGAT 649
QY 712 CCAGCGCCACACTTTGGGCGTGACAGCGCGGTGACATGCTGCGCGAAACCCAC 771
DB 648 CCAGCGCGACACTTTGGGCGTGAGCAGCGGCGTGACATGCTGCGCGAAACCCAC 589
QY 772 CAGCGGGAACAGATGAGGTTGCGCGAACCGCGCGCGGGAATTCACCGTCGGT 831
DB 588 CAGCGGGAACAGATGAGGTTGCGCGAACCGCGCGCGGGAATTCACCGTCGGT 529
QY 832 GTGCGGTTATGCGGCGGCGCGCGCGGATTAACACAGCGGTCCAATACGANTCG 891
DB 528 GTGCGGTTATGCGGCGGCGCGCGCGGATTAACACAGCGGTCCAATACGANTCG 469
QY 892 GTTCAACAGGTGACGACGATGCGCATCAATGATGCGTTTACCAACCGCGCTTATC 951
DB 468 GTTCAACAGGTGACGACGATGCGCATCAATGATGCGTTTACCAACCGCGCTTATC 409
QY 952 GAGCGTCACCTCCACCGCGACCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1011
DB 408 GAGCGTCACCTCCACCGCGACCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 349
QY 1012 AGCGGTGTCATCGGCGAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1071
DB 348 AGCGGTGTCATCGGCGAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 289
QY 1072 GTTGTGTCGCGGTCTTTGCGTCCATGACGCGCAACCAATACAGGTGAAGCGATGCC 1131
DB 288 GTTGTGTCGCGGTCTTTGCGTCCATGACGCGCAACCAATACAGGTGAAGCGATGCC 229
QY 1132 CCAGCGGATAATATGACGACGATCGCGCGGCGATTTGGACAAAGATCAACGCGCGAAGT 1191
DB 228 CCAGCGGATAATATGACGACGATCGCGCGGCGATTTGGACAAAGATCAACGCGCGAAGT 169
QY 1192 GCCGCGGATGAACAAAGAGGTGCAAAATTAACACAGAGAGAACCGCAATGATGCC 1251
DB 168 GCCGCGGATGAACAAAGAGGTGCAAAATTAACACAGAGAGAACCGCAATGATGCC 109
QY 1252 TTCGCGGTTAAATTCCTTTTAAATCACAGTACATTTCTGCGGTTCGATGAGCAGTAAAG 1311
DB 108 TTCGCGGTTAAATTCCTTTTAAATCACAGTACATTTCTGCGGTTCGATGAGCAGTAAAG 49

QY 1312 ACTGGCCCCCAAGAGCAGACCTGTATGAGATTTCATGATCACCAT 1359
DB 48 ACTGGCCCCCAAGAGCAGACCTGTATGAGATTTCATGATCACCAT 1
RESULT 11
AAH45375/c
ID AAH45375 standard; DNA; 1568 BP.
XX AAH45375;
AC
XX 11-SEP-2001 (first entry)
DT
XX C. thermoaminogenes lysin biosynthetic enzyme lyse DNA.
DE
XX Heat-resistant; lysin biosynthesis; enzyme; coryneform;
KW aspartate-semialdehyde dehydrogenase; lyse; ds.
KW
OS Corynebacterium thermoaminogenes.
XX JP2001120270-A.
PN
XX 08-MAY-2001.
PD
XX 01-NOV-1999; 99JP-00311148.
PF
XX 01-NOV-1999; 99JP-00311148.
PR
XX (AJIN) AJINOMOTO KK.
PA
XX WPI; 2001-364760/38.
DR P-PSDB; AAG64047.
DD
XX A heat-resistant lysin biosynthetic system enzyme gene of a high
PT temperature-resistant coryneform microbe.
PT
XX Example 5; Page 22-24; 27pp; Japanese.
FS
XX The invention relates to a gene from a high temperature-resistant
CC coryneform microbe that encodes a heat-resistant lysin biosynthetic
CC enzyme. The enzyme has aspartate-semialdehyde dehydrogenase activity and
CC can be used for growing amino acid-producing microbes. The present
CC sequence encodes an enzyme of the invention
XX
SQ Sequence 1568 BP; 258 A; 525 C; 490 G; 295 T; 0 U; 0 Other;
Query Match 29.4%; Score 698.6; DB 4; Length 1568;
Best Local Similarity 68.4%; Pred. No. 7e-207;
Matches 1017; Conservative 0; Mismatches 454; Indels 16; Gaps 3;
QY 650 TAACCCATCAACATCAGTTTGTATGGCCAAATGCGGTATCACTGCGACGACGCTTG 709
DB 1472 TAACCCATCAGGATCAGCTTCACGGCCAAATCGGTGAGCACCACGCGCCATATGTTG 1413
QY 710 ATCCAGCGCCACACCTTGGGCGTGGACGCGGCGTGCATGCTGCGCGCAACCC 769
DB 1412 ATCCAGCGCCACACCTTGGGCGTGGACGCGGCGTGGACGCGGCGTGGACGCGG 1353
QY 770 ACCAGCGGGAACAGATCAGGCTTGGCGGAACCGCGCGCGGCAAAATCCACCGTCCG 829
DB 1352 ACCAGTGGGAACAGATCAGGCTTGGCGGCGGCGGCGGCGGCAACCCAGCGACCG 1293
QY 830 GTGTGCGCGGTATTTGGCGCGCGACGCGCGCGATTAACACAAACGCGTCCAAATAGCGATTC 889
DB 1292 GTCTCCCGGCTACTGGGCTCCGACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1233
QY 890 GGGTTCAACAGGTGACGCGGATTTGCCATCACTGCGGCTTTACCCCAACCGCTGCTTA 949
DB 1232 GGAATTGAGCGAGTCAACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1173
QY 950 TCGACGCTCACTCCACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1009
DB 1172 CCTGATGTGATGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1113

QY	1010	AAAGGGGTGTCATCGGGCAGGGTTGGTTCTGTGTTCTTCAATGATCTGTGGCGCTTCCACC	1069
DB	1112	GCTGCCG---CAACGGGTTCCGAATGCTC-----GACAAAGGTTTACCTCGGTG	1068
QY	1070	TTGTTTGTCTAGCGTCTTTTCGGTCCCATGACGGCAACCATAAACAGTAAGCGATGCCA	1129
DB	1067	CGSGCGCGCAGGGCGTCCGGCGCGCCATCACCGGGAACACAGCAGGTAGGCGATGCG	1008
QY	1130	CCCCAGCGCATATATCTAGACGACGATCGGCGGGCATTTGGACAAAAGATCAACGCCCAAG	1189
DB	1007	CACACGCGCAGGATGTGAGAAATGATCGGGCGGTGTGCGAGATCAGGCCGACCCCGAGG	948
QY	1190	GTGCGGGGATGACAAAAGACCTCAGAAATTAAACACACGAGAAAGACCGCATGAGT	1249
DB	947	GTGCGGAGGTGAACAGGACCCGCTGGGACACGACAGACAGCATGATGACGGCGGTGATG	888
QY	1250	CTTCCGCGCTTAATTCCTTTGTTTAATCCACAGTACATTTCTGGCGTCCCGATGGACGATAA	1309
DB	887	CCCTCGCGTTTGATGCCCTGTGTGATCACCAGGACATCTGTGGCGCGATGGCCACAGC	828
QY	1310	AGACTGGCCCCAAAAGCAGACCTGTATGAAGATTTCCATGATCAACATCGATCGTGACAT	1369
DB	827	AGACTGGCTCCCAACACAAACCGGTGACAAAGATTTCCATGTCCCGCATGTTCCGCGAC	768
QY	1370	GGAGTACTTAAGTAAATGATGGTTCTTACATGTTTTATATAGCTTCATGAACCCC	1429
DB	767	AAACCACATTAAGTAAATGATGAGACTGAACATGTTTTAGAAATTGCTTCATGAACCCG	708
QY	1430	ATTCAACTGGACACTTTGCTCTCAATCATTTGATGAAGCGAGCTTCGAAGCGCTCCTTTA	1489
DB	707	ATCCACCTGGACACCTTCTGACCATCATCGACGAGGCGAGCTTCGAGACGCTCCCTC	648
QY	1490	GCCCTTTCATTTCCCTTCGGGGTGAGTGAAGCGTTTAAAGTCTTCGAGACATCACGTG	1549
DB	647	GCACTGTGCATCTCCCTTCGGCGGTGACGAGCGCATCAAGGCACACTGGAGAAATCCGTC	588
QY	1550	GGTCCAGTGTGTGATTCGCGCACCCACCGGCCAAGACACGAGCGGGTGAAGTCCTT	1609
DB	587	GGTCCGGTACTGTGTGCGGACCCACGCGCGGTTCGCACGAAGCCGGGAGTGTG	528
QY	1610	GTGCAAGCAGCGCGGAAATGGTTGCTGCAAGCAGAACTAAAGCGCAACTATCTGGA	1669
DB	527	GTCCAGCGCGCGCAGATGGCGTCTGCGAGCGGAGACCGGTGACCACTGGCGAA	468
QY	1670	CGCTTGTCTGAATTCGGTTTAACTATGCGCATCAACGCGATTCGCTATCAATGATGTTT	1729
DB	467	CGGCTTGAAGAGATCCCCCTGACCGGTGGGTATCAACCGCGATTTCCCTGTCCACTGGTTC	408
QY	1730	CTTCCCGTGTTCAAAGAGTAGTCTTTCGGGGTGGACACGCTCAGCGTGGCTTGGAA	1789
DB	407	CGCCCGTGTTCGCCAGGTGTGCACATTTGGGGTGGGTCAACCTCGCTGGGTGTGGAG	348
QY	1790	GATGAAGCGCACATTAATCTCTGCTCGCGCGTGGAGATGTTTTAGGAGCGGTAAACCGGT	1849
DB	347	GACGAGGCCACACCTGTCTCTGTTGCGCGCGGTTCCGTGCTCGGGGCGGTGACCCGC	288
QY	1850	GAAGCTAATTCCTGGCGGGGATGTGAAGTAGTAGAACTTGGAAACATGGCGCACATTTGGCC	1909
DB	287	GAGGCGGACCCCGTTGCGCGGTGTGAGGTTCCTAGCGCTCGGGGTCAATGAGACACTGCGCG	228
QY	1910	ATTGCAACCCCTCATTTGGGGATGCGCTACTAGTGTGATGGGAACACTAGATTGGGCTCG	1969
DB	227	GTGGCCACCCCGGAATCTGAGGACACGGGTACACCGGTGATGGCCACCCCGACTTGGGTGCG	168
QY	1970	-ATGGCCCGTCTTACGCTTCGGTCCCAAGATGTGCTTCAAGACCTGTGACCTGACGGCGC	2028
DB	167	CATGCGGCTACTGGGTTTCGGCCCGGAATGATGTCTCAGBACCGGATCTCGAGGGCAG	108
QY	2029	CGTCAATGGTCTGTGGGGCGAGGCGGTATTCATTTGTTCCTCGCGGAGAGGTTTTTGG	2088
DB	107	GGTCAACGAGACCGGTTCGACAGCGCGCTGTCTCGGTTGGTGGCTGTGCGAGGGTTTCGG	48

Qy	2089	TGAGGGCAATTCCCGGACGGCCTTGCTTGGGACACTTTCTTCGCCGAACCC	2133
Dd	47	GGAGGCGGTGCAGACTCGCTCTGGGGTGGGGCCCTGCTGCCCCGAGGCC	1
 RESULT 12 ACC80942/c			
ID	ACC80942 standard; DNA; 712 BP.		
XX AC	ACC80942;		
XX DT	27-OCT-2003 (revised)		
DT	11-AUG-2003 (first entry)		
XX XX	LysE24 protein encoding sequence.		
DE			
XX			
KW	L-lysine; L-arginine; LysE24; ds.		
XX OS	Corynebacterium glutamicum.		
PH	Key Location/Qualifiers		
FT	1..711		
FT	/tag= a		
FT	/product= "LysE protein"		
PX	BP1266966-A2.		
PN			
PD	18-DEC-2002.		
XX			
PF	05-JUN-2002; 2002EP-00012539.		
XX			
PR	12-JUN-2001; 2001JP-00177075.		
PA	(AJIN) AJINOMOTO CO INC.		
PI	Gunji Y, Yasueda H;		
DR	WPI; 2003-241171/24.		
XX	F-PSDB; ABR58214.		
PT	Novel DNA encoding variant of LysE protein from a coryneform b		
PT	when introduced into methanol assimilating bacterium, facilitat		
PT	excretion of L-lysine and/or L-arginine to outside of a cell.		
XX PS	Example 1; Page 19-20; 23pp; English.		
XX CC	The present invention relates to DNA encoding variants of prot		
CC	loop region and six hydrophobic helices which facilitates exc		
CC	lysine and/or L-arginine to outside of cell of a methanol assi		
CC	bacterium when introduced into the bacterium. The method is us		
CC	encoding a protein which facilitates excretion of L-lysine, L-		
CC	both of these L-amino acids to outside of a cell of a methano		
CC	assimilating bacterium when DNA of the method is introduced i		
CC	bacterium. The present sequence represents a LysE24 protein fr		
CC	Brevibacterium lactofermentum encoding sequence. (Updated on 2		
CC	to standardise OS field)		
XX			
SQ	Sequence 712 BP; 133 A; 173 C; 224 G; 182 T; 0 U; 0 Other;		
 Query Match 29.4%; Score 696.8; DB 8; Length 712; Best Local Similarity 99.6%; Pred. No. 1.6e-206; Matches 709; Conservative 0; Mismatches 2; Indels 1			
Qy	649	CTAACCCATCAATCATGTTGATGGCCAATGGGTGATCATCAAACTGCCACGACG	
Dd	712	CTAACCCATCAATCATGTTGATGGCCAATGGGTGATCATCAAACTGCCACGACG	
Qy	709	GATCCAGCGCCACACTTTGGGGCTTGACAGCGGGGTGACAATGCTGCTGTCGCG	
Dd	652	GATCCAGCGCCACACTTTGGGGCTTGACAGCGGGGTGACAATGCTGCTGTCGCG	
Qy	769	CACACAGCGGACCAAGATCATGGCTTGGCCGCGAACCGCGCACGCGCGGAAAATCG	

AC AAF68077;
 DT 11-APR-2001 (first entry)
 DE Corynebacterium glutamicum MCT protein encoding DNA SEQ ID NO:669.
 XX Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
 KW membrane construction and membrane transport protein; petroleum spill;
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
 KW identification; microorganism; fine chemical production; transformation;
 KW genome mapping; genetic engineering; ds.
 XX Corynebacterium glutamicum.
 OS
 XX
 PN WO200100805-A2.
 XX 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-1800926.
 XX
 PR 25-JUN-1999; 99US-0141031P.
 PR 08-JUL-1999; 99DE-01031454.
 PR 08-JUL-1999; 99DE-01031478.
 PR 08-JUL-1999; 99DE-01031563.
 PR 09-JUL-1999; 99DE-01032122.
 PR 09-JUL-1999; 99DE-01032124.
 PR 09-JUL-1999; 99DE-01032125.
 PR 09-JUL-1999; 99DE-01032128.
 PR 09-JUL-1999; 99DE-01032180.
 PR 09-JUL-1999; 99DE-01032182.
 PR 09-JUL-1999; 99DE-01032190.
 PR 09-JUL-1999; 99DE-01032191.
 PR 09-JUL-1999; 99DE-01032209.
 PR 09-JUL-1999; 99DE-01032212.
 PR 09-JUL-1999; 99DE-01032227.
 PR 09-JUL-1999; 99DE-01032228.
 PR 09-JUL-1999; 99DE-01032229.
 PR 09-JUL-1999; 99DE-01032300.
 PR 14-JUL-1999; 99DE-01032327.
 PR 14-JUL-1999; 99DE-01033005.
 PR 14-JUL-1999; 99DE-01033006.
 PR 27-AUG-1999; 99DE-01040764.
 PR 27-AUG-1999; 99DE-01040765.
 PR 27-AUG-1999; 99DE-01040766.
 PR 27-AUG-1999; 99DE-01040830.
 PR 27-AUG-1999; 99DE-01040831.
 PR 27-AUG-1999; 99DE-01040832.
 PR 27-AUG-1999; 99DE-01040833.
 PR 31-AUG-1999; 99DE-01041378.
 PR 31-AUG-1999; 99DE-01041379.
 PR 31-AUG-1999; 99DE-01041395.
 PR 03-SEP-1999; 99DE-01042077.
 PR 03-SEP-1999; 99DE-01042078.
 PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042088.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 DR WPI; 2001-071486/08.
 DR P-PSDB; AAB76844.
 XX
 PT Corynebacterium glutamicum nucleic acids encoding membrane construction
 PT and membrane transport proteins or their portions, useful for typing or
 PT identifying C. glutamicum or related bacteria, and as markers for
 PT transformation.
 XX
 PS Claim 3; Page 1107-1108; 1119pp; English.
 XX
 CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
 CC construction and membrane transport (MCT) proteins given in AAB76510 to
 CC AAB76847. The MCT nucleic acids and proteins are useful in the

CC identification of microorganisms which can be used to produce fine
 CC chemicals, for modulating fine chemical production in C. glutamicum or
 CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
 CC identification of C. glutamicum or related bacteria, as reference points
 CC for mapping C. glutamicum genome, and as markers for transformation.
 CC AAF68082 and AAF68082 represent sequencing primers which are used in an
 CC example from the present invention
 XX
 SQ Sequence 993 BP; 222 A; 247 C; 277 G; 247 T; 0 U; 0 Other;
 Query Match 20.4%; Score 485.2; DB 4; Length 993;
 Best Local Similarity 99.0%; Pred. No. 2.7e-140;
 Matches 499; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
 QY 1 AGATACTCTTTGGAGAAACCATGTACGATTGGCTGACATTGTCGTCTGAAAGGC 60
 DB 490 AGATACTCTTTGGAGAAACCATGTACGATTGGCTGACATTGTCGTCTGAAAGGC 549
 QY 61 TCTTTACGTGGGTATTTCTTCTTCTACGCTCCAGAGCTCACAGCGAGGCGGCTGAGTTCA 120
 DB 550 TCTTTACGTGGGTATTTCTTCTTCTACGCTCCAGAGCTCACAGCGAGGCGGCTGAGTTCA 609
 QY 121 GCGGAGGAGGGCTGCGCGGCTTCTGATTCATCAGCAAGCTATTCATCATTAATCGTTG 180
 DB 610 GCGGAGGAGGGCTGCGCGGCTTCTGATTCATCAGCAAGCTATTCATCATTAATCGTTG 669
 QY 181 GGTGGAGGAACCGGGCGATGACGCTGAGAACTTTGTCAGTCAGCTGCAACAATGGTCT 240
 DB 670 GGTGGAGGAACCGGGCGATGACGCTGAGAACTTTGTCAGTCAGCTGCAACAATGGTCT 729
 QY 241 TGGCGTCAATGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 300
 DB 730 TGGCGTCAATGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 789
 QY 301 AATTCCAGAGGGTTCCCGCGCCAGCCAGGCTTAAGTCCCTGCTGAGGCGCATGTTGAAGT 360
 DB 790 AATTCCAGAGGGTTCCCGCGCCAGCCAGGCTTAAGTCCCTGCTGAGGCGCATGTTGAAGT 849
 QY 361 GAACAATATTGATATGCTTCCGCAAGCTCAATGACATCGCCAGGAAACGCGGCGAGTCACT 420
 DB 850 GAACAATATTGATATGCTTCCGCAAGCTCAATGACATCGCCAGGAAACGCGGCGAGTCACT 909
 QY 421 TGGCGAGATGCGGCTTGCATGCGGCTGCGGAGCAAGGAGAGTACGCGCGC--GGATACC 478
 DB 910 TGGCGAGATGCGGCTTGCATGCGGCTGCGGAGCAAGGAGAGTACGCGCGCGCGGATTACC 969
 QY 479 GTGACCACTGCAATGATGCTGCT 502
 DB 970 GTGACCACTGCAATGATGCTGCT 993
 RESULT 15
 ACA29651
 ID ACA29651 standard; DNA, 1095 BP.
 XX
 AC ACA29651;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #11308.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Corynebacterium diphtheriae.
 XX
 PN WO200271183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.

Search completed: March 15, 2004, 14:59:19
Job time : 971.046 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:47:35 ; Search time 178.837 Seconds
(without alignments)
7366.769 Million cell updates/sec

Title: US-09-105-117K-3
Perfect score: 2374
Sequence: 1 agatactcttggagaaga.....gtacacattcagcaaatgg 2374

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/prodata/2/ina/5A COMB.seq.*
2: /cgm2_6/prodata/2/ina/5B COMB.seq.*
3: /cgm2_6/prodata/2/ina/6A COMB.seq.*
4: /cgm2_6/prodata/2/ina/6B COMB.seq.*
5: /cgm2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgm2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122.8	5.2	15239	1 US-08-390-878-17	Sequence 17, Appl
2	121.2	5.1	909	4 US-09-894-844-12	Sequence 12, Appl
3	121.2	5.1	4403765	3 US-09-103-840A-2	Sequence 2, Appl
4	121.2	5.1	4411529	3 US-09-103-840A-1	Sequence 1, Appl
5	100.8	4.2	5541	1 US-08-920-812-20	Sequence 20, Appl
6	100.8	4.2	5541	1 US-08-920-812-20	Sequence 20, Appl
7	100.8	4.2	5541	1 US-08-921-177-20	Sequence 20, Appl
8	100.8	4.2	5541	1 US-08-362-577C-20	Sequence 20, Appl
9	100.8	4.2	5541	2 US-08-920-828-20	Sequence 20, Appl
10	81.4	3.4	1095	4 US-09-489-039A-5370	Sequence 5370, Ap
11	80.2	3.4	915	4 US-09-489-039A-4735	Sequence 4735, Ap
12	73.4	3.1	936	4 US-09-543-681A-791	Sequence 791, Ap
13	65.2	2.7	900	4 US-09-252-991A-7644	Sequence 7644, Ap
14	64	2.7	1002	4 US-09-724-623-24	Sequence 24, Appl
15	61.8	2.6	699	4 US-09-252-991A-7908	Sequence 7908, Ap
16	61.8	2.6	834	4 US-09-252-991A-7643	Sequence 7643, Ap
17	61.8	2.6	894	4 US-09-252-991A-7841	Sequence 7841, Ap
18	57	2.4	597	4 US-09-894-844-11	Sequence 11, Appl
19	56.6	2.4	642	4 US-09-489-039A-4674	Sequence 4674, Ap
20	54.6	2.3	903	4 US-09-328-352-1694	Sequence 1694, Ap
21	53.8	2.3	1176	4 US-09-252-991A-7571	Sequence 7571, Ap
22	49.2	2.1	921	4 US-09-252-991A-14393	Sequence 14393, A
23	49.2	2.1	1905	4 US-09-252-991A-14312	Sequence 14312, A
24	47.8	2.0	474	4 US-09-252-991A-7572	Sequence 7572, Ap
25	46.2	1.9	903	4 US-09-252-991A-5488	Sequence 5488, Ap
26	45.8	1.9	720	4 US-09-252-991A-7723	Sequence 7723, Ap
27	45.4	1.9	45613	4 US-09-596-002-22	Sequence 22, Appl

c	28	43	1.8	645	4	US-09-543-681A-800	Sequence 800, App
c	29	42.8	1.8	1879	4	US-09-453-702B-142	Sequence 142, App
c	30	42	1.8	936	4	US-09-252-991A-16286	Sequence 16286, A
c	31	42	1.8	978	4	US-09-252-991A-15990	Sequence 15990, A
c	32	42	1.8	1533	4	US-09-252-991A-16513	Sequence 16513, A
c	33	41.6	1.8	505	4	US-09-621-976-15639	Sequence 15639, A
c	34	40.6	1.7	969	4	US-09-252-991A-345	Sequence 345, App
c	35	40.4	1.7	285	4	US-09-252-991A-5423	Sequence 5423, App
c	36	40	1.7	864	4	US-09-252-991A-12026	Sequence 12026, A
c	37	40	1.7	870	4	US-09-252-991A-11957	Sequence 11957, A
c	38	39.6	1.7	1935	2	US-08-492-027A-9	Sequence 9, Appli
c	39	39.4	1.7	505	4	US-09-621-976-15639	Sequence 15639, A
c	40	39.4	1.7	927	4	US-09-252-991A-12324	Sequence 12324, A
c	41	39.4	1.7	972	4	US-09-252-991A-12096	Sequence 12096, A
c	42	39.4	1.7	1482	4	US-09-252-991A-12421	Sequence 12421, A
c	43	38.6	1.6	924	4	US-09-252-991A-12055	Sequence 12055, A
c	44	37.6	1.6	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	45	37.6	1.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-390-878-17
; Sequence 17, Application US/08390878
; Patent No. 5700683
; GENERAL INFORMATION:
; APPLICANT: Stover, Charles K.
; APPLICANT: Mahairas, Gregory G.
; TITLE OF INVENTION: VIRULENCE-ATTENUATING GENETIC DELETIONS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower, 20th
; STREET: Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,878
; FILING DATE: 17-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 15371A-17
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/543/9600
; TELEFAX: 415/543/5043
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15239 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-390-878-17

Query Match 5.2%; Score 122.8; DB 1; Length 15239;
Best Local Similarity 49.7%; Pred. No. 1.1e-25;
Matches 434; Conservative 0; Mismatches 422; Indels 18; Gaps 4;

QY 1433 CAACCTGGACATTGGCTCTCAATCATGTGAGGAGGAGCTTCGAAGCGGCTCTTAGCC 1492
DB 5253 CAGCTGGCCGATTTGGTCCCTGGTCTGACTGGGAGCTTGATGGCGCGGAGGCG 5312

Db 2229522 GAGGACGAGGACATTCCGCGCGGTCTACGAGGGGTGTGGCGATGGGCGGTGACC 2229463
 Qy 1847 CGTGAAGCTAATCCGCGGCGGAGTGAAGTGTAGAACTTTGGAACTGGGCGGCACTTG 1906
 Db 2229462 ACCGAGCGGAACCGGCGGCGGTGCGCGGTGACCCGCTGGGTGAATGCGCTACCTA 2229403
 Qy 1907 GCATTTGCAACCCCTCATTTGGGATGCTTACATGGTGTGATGGGAACTAGATTGGCT 1966
 Db 2229402 CCAGTGGCAGCAGCCATTCGTCAGCGCCATCTATCCGAGGGGTTCATTCGCCCGCGG 2229343
 Qy 1967 GCGATGCCGCTTACGCTTCGCTCCAAAGATGTCTTCAAGACCGGTGACCTGGACGGG 2026
 Db 2229342 GCGCTAAGCTCCGCTCACTGGCGTGGATCGTCAAGTGGGCTCGAGCATGTTGGTG 2229283
 Qy 2027 GCG---GTGATGTCTGTGGGGGCGAGGCGGTATCTCATTTGTCCTGGGCGGAGGT 2093
 Db 2229282 CGTAAGGCGCTTTCGTCGCGCATCACAGACCGAGCAGCTTGTCTCCGACACAGAGGC 2229223
 Qy 2084 TTTGGTGAGGCAATTCGCGAGGCTTGTGGGATCTTCTCCGAAACCAAGCTGCT 2143
 Db 2229222 TTACGCGCGAGCGCGCGGCTGGGATGGGCGATGTTCCCGAGAGCTGGAGCA 2229163
 Qy 2144 CCCATGCTAAAGCAGGAGAGTGTATCTCTCGATGAGATACCCATTTGACACACCGATG 2203
 Db 2229162 TCTCCGCTTCCGATGATGCTTGTGTCGAGGGTCTGGACATACCTCGAGCTCCCTTC 2229103
 Qy 2204 TATTTGCAAGTGGCGCTTGAATCTAGATCTCTAGTACTCAGAGCGCGCTGTT 2263
 Db 2229102 TATTTGCAATGCTGGAATGAGACAGTCCGATCATCCGCGAATACCGACAGCGTGAGG 2229043
 Qy 2264 GATGCAAGCAATCGAGGATTCGCGCTTAGTTAC 22297
 Db 2229042 GCGCGGCAAGCGGTCTGTACCGGGGCGGCAAC 2229009

RESULT 5

US-08-920-812-20/c
 ; Sequence 20, Application US/08920812
 ; Patent No. 5763188
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohno, Tsuneya
 ; APPLICANT: Matsuhisa, Akio
 ; APPLICANT: Uehara, Hirotsugu
 ; APPLICANT: Eda, Soji
 ; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/920,812
 ; FILING DATE: 29-AUG-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/362,577
 ; FILING DATE: 27-MAR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures, Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 19036/32420
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5541 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Genomic DNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Escherichia coli
 ; STRAIN: Clinical Isolate EC-625
 ; US-08-920-812-20
 Query Match 4.2%; Score 100.8; DB 1; Length 5541;
 Best Local Similarity 50.5%; Pred. No. 2.2e-19;
 Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;
 Qy 1 AGATPACTCTTTTGAAGAACAACATGTACGCATGCGGTGACATTTGTGGTCTTGAAGGC 60
 Db 3416 AATATCGCCGATGGAAGAAACCCGCTCTCGCTGGCTCATGCGGTACAAAGCGGTAAAGC 3357
 Qy 61 TCTTTAGTGGGTATTTCTTCTACGGTCCAGAGCTCAGCGGAGGCGGTGATTCAT 120
 Db 3356 GCTTTATGTGCGGATCTCTCTTCTACTCGCCAGAGCGGACGCAAAAATGTTGATGTCT 3297
 Qy 121 GCGCGAGGAGGCTCCCGCTTCTGATTCATCAGCAAGCTATTTCCATCATTAATCGTTG 180
 Db 3296 GCGGAGTGGAAATTCGCTGTGTTAATTCATCAACCTTGTACAAATTACTGAACCGCTG 3237
 Qy 181 GGTGAGGAACCGGCGGATGAGCGGTGAGAACTTTGTGAGTCTGAGTCCCAACATGTCT 240
 Db 3236 GGTGGATAAAGCGG-----CCTGTGATACCTTCGCAAAATAACGCGCT 3192
 Qy 241 TGGCGTCAATGCTTCTCACCAGTTCGCGAGGCTGCTCAGGACAAATATCTCGATGG 300
 Db 3191 GGGCTGTATTTGCTTACTCTCTGGCTCAGGATTTCTGACCGGAATATCTCAACGG 3132
 Qy 301 AATTCCAGAGGTTCCCG-----CCGAGCGAGGTAAGTCCCTGTCTGAGGG 348
 Db 3131 CATTCGCAAGATTCACGATGATCGTGAAGGGAATAAAGTTCTGTTGCTGACCGCGAA 3072
 Qy 349 CATGTTGAAGTGAACATATTGATGTCTCGAGCTCAATGATCATCGCCAGGACG 408
 Db 3071 AATGCTTACGAAAGCCAACTCAACAGCCTTACCTTATTGAATGAATGGCAAGAGCG 3012
 Qy 409 CGGCGAGTCACTTGGCGAGATGGCGCTTGCATGGGTGCTGCGGAGCAAGAGAGTACGG 468
 Db 3011 TGGACATCAATGCGGCAATGGCGTTAAGCTGTTGCTGAAGA-----2967
 Qy 469 CGCGGATACCGTGAACAGTGTATGATGTTGTTGCTTCTGCTGAGTTCAGCAGTGGACAG 528
 Db 2966 --TGATCGGTGAGCTCGGTATTGATTGGTGC-CAGCGCGCGGACAACTTTGAGGAGAA 2911
 Qy 529 CTTTGATTCACTCAAACTTTGGAGTTTCTGACCGCGAGTTGGAGCGATCGATGAGAT 588
 Db 2910 CTTGAGCGCTGATTAATCTGACATTTAGCACCGAGGAGCTGGCCAGATTGATCAGCA 2851
 Qy 589 TTCCCGACGACCGCGCATCAACATTTGGGCGAAGCCACCGATTTCCAAACCCCGGAAAA 648
 Db 2850 TATCGCGATGGCGAGCTGAATCTGTGCGAGCGGTCTTCCGATAAATGACCTGTATAAA 2791
 Qy 649 CTAAACCATCAACATCAGTTTGTATGGCCATTCGGGTATCACTCAACTGCCACGACGATT 708
 Db 2790 CGGCGCGAGAACCGCGCGCCGATTAAGATCAGTGAATATGACGAGTCAAGTCCGATCAAGAT 2731
 Qy 709 GATCCAGGCGCACACCTTTGGGCTGGACAGCGCGGCGTGACAAATG 752
 Db 2730 AGCCCATCAAAATGACAGAGACCAAAAGTCAAGTGAATATG 2687

RESULT 6

US-08-920-827-20/c
 ; Sequence 20, Application US/08920827
 ; Patent No. 5770375


```

;
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: Clinical Isolate EC-625
; US-08-920-827-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;
Best Local Similarity 50.5%; Pred. No. 2.2e-19;
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

QY 1 AGATACCTCTTTGGAGAAACCATGTACGCAITGGGTGACATTTGTCGCTCTGGAAGGC 60
DB 3416 AAATACCGCATGGAAGAACCGCTCTGGCTGCTCATGCGGTACAAAGCGGTAAAGC 3357
QY 61 TCTTTAGTGGTATTTCTTCTACGTCCAGAGCTCAGCGGAGCGGCTGATTCAT 120
DB 3356 GCTTTATGTCCGATCTCTCTTACTCTCCGAGAGCGGACGCAAAATGTTGATGCT 3297
QY 121 GCGGAGAGGCGTCCGCTCTTCAATTCATCAGCAAGCTATTCCATCATTAATCGTTG 180
DB 3296 GCGGAGTGAATAATTCGCTGTTAATTCATCACTTCGTCATTAATTCGACCGCTG 3237
QY 181 GGTGAGGAACCGGCGATGACGGTGAGAACTTTGTCAGTCAGCTGCCAATGCTCT 240
DB 3236 GGTGATAAAGCGG-----CCTGCTGATACCTCGCAAAATGACGGCT 3192
QY 241 TGGGCTCATTTGCTTTCTACACATTCGCGAGGCGCTGCTCAGGCAATATCTGATGG 300
DB 3191 GGGCTGATTTGCTTTTACTCTCTGCTCAGGAGTTCGTCACCGGAAATATCTCAACGG 3132
QY 301 AATTCAGAGGGTTTCCG-----CGCAGCCAGGGTAAGTCCCTCTCTGAGGG 348
DB 3131 CATTCGCAAGATTCAGGATGCAATCGTAGGGAATAAAGTTCGTGGTCTGACCCGAA 3072

```

```

QY 349 CATGTTGAAGTGAACAAATATTGATATGCTCCGACAGCTCAATGACATCCGCCAGCAAG 408
DB 3071 AATGCTTTACCAAGCCCAACCTCAACAGCTTACACTATTGTAATGAATGGCAGCAGCG 3012
QY 409 CGGGCAGTCACTTGGCGAGATGGCGTTTGCATGGGTGCTGCGCGAGCAAGGAGATACGG 468
DB 3011 TGGACAATCAATGGCGCAATGGCGTTAAGCTGTTGCTGAAAGA-----2967
QY 469 CGGGATACCTGACCAAGTGCATTTGATTTGGTCTTCTGTCAGTTGAGCAGCTGGAACAAG 528
DB 2966 ---TGATCGCTGACGTCGCTGATTTGATTTGGTGC-CAGCGCGCGGCAACTTGAGGAGAA 2911
QY 529 CTTTGAATCACTCAACAACTTGGAGTTTCTGACGCGGAGTTGGAGCGATCGATGAGAT 588
DB 2910 CGTGACGGCTGATTAATCTGACATTTAGCAGGAGGAGCTGGCGCAGATTGATCAGA 2851
QY 589 TTCCACGACCGCGCATCAACATTTGGGCGAAGGCCACCGATTCAAAAACCGCGAAAA 648
DB 2850 TATCGCGCATGGCGAGCTGAATCTGTGGCAGCGCTCTCCGATAAATGACCTGTTAATA 2791
QY 649 CTAACCCATCAACATCAGTTTGTATGGCCATGGGTTCATCACAACCTGCCACGACGCTT 708
DB 2790 CGGGCGGAGAACCGCGCCCGCATTAAGATCAGTGATTATGACAGTCAGTCGCTCAAGAT 2731
QY 709 GATCCAGCGCCACACCTTGGGCTGGACAGCGGCGCTGACAAATG 752
DB 2730 AGCCCATCAAAATGACAGAGACACAACTGATGATGATG 2687

RESULT 7
US-08-921-177-20/c
; Sequence 20, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: Clinical Isolate EC-625
; US-08-921-177-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;
Best Local Similarity 50.5%; Pred. No. 2.2e-19;
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

QY 1 AGATATCTCTTTGGAAGAACCATGACGATTCGCTGACATTTGCGTCTGGAAGGC 60
DB 3416 AATATACCGCATGGAAGAACCGCTCTGCGTGGCTCATGCGGTACAAAGCGGTAAGC 3357

QY 61 TCCTTACGTGGTATTTCTTCTACGGTCCAGAGCTCAGCGAGGCGGTGATGAT 120
DB 3356 GCTTTATGTGCGGATCTCTCTTACTCGCAGAGCGGCAAAAATGTTGATGTCT 3297

QY 121 GCGGAGAGGCGTCCCGCTCTGATTCTATCAGCAAGCTATTCATCATTAATCGTTG 180
DB 3296 GCGGAGTGGAAAATTCGCTGTTTAACTTCACTTCTGTAACAATTTACTGAACCGCTG 3237

QY 181 GGTGAGAACCGCGGATGACGTGAGAACCTTGTGACGTGAGTCCCAATGCTCT 240
DB 3236 GGTGAGTAAAGCGG-----CCTGCGATACCTTCGCAAAATAACGCGCT 3192

QY 241 TGGGCTCATTTCTTCTCACCCTTGGCAGGCGCTGCTCAGCGACAAATATCTCGATGG 300
DB 3191 GGGCTGATTTGCTTCTTCTCTGCGTCAAGGATTTGCTGACCGAATAATCTCAACGG 3132

QY 301 AATTCAGAGGCTTCCCG-----CGCCAGCGAGGTAAGTCCCTGTCTGAGGG 348
DB 3131 CATTCGCAAGATTCACGATGATCGTGAAGGAATAAAGTTCTGTTGCTGACCGGAA 3072

QY 349 CATGTTGAACGTGAACATATTTGATGTTGTCGCAAGCTCAATGACATCGCCAGGAACG 408
DB 3071 AATGCTTACCGAAGCAACCTCAACAGCCTACATTTATTGAATGAATGGCACAGACGG 3012

QY 409 CGGCGAGTCACTTTGGCGAGATGGCGCTTTCATGGGTGCTGCGGAGCAAGGAGTACGG 468
DB 3011 TGGCAATCAATGGCGCAATGGCGTTAAGCTGTTGCTGAAAGA-----2967

QY 469 CGCGATACCGTGACGATGCAATGATTTGTTGCTTCTGATGAGAGCTGGAACAACG 528
DB 2966 ---TGATCGGTGACGTCGGTATTTGATTTGTTGTC-CAGCGCGCGGACACTTGAGNGAA 2911

QY 529 CTTGATTCATCAAACTTTGAGATTTTCTGACCGCGAGTTGGAGGCGATCGATGAT 588
DB 2910 CGTGAGGCGGCTGAATATCTGACATTTAGCACCGAGGAGCTGGCGAGATTGATCAGCA 2851

QY 589 TTCCGACGCGCGCATCAACATTTGGCGAGGCGCCAGCTATTCAAAACCGCGAAA 648
DB 2850 TATCGCGATGGGAGCTGAATCTGTGGCAGCGCTCTTCCGATTAATGACCTGTTATTA 2791

QY 649 CTAACCCATCAACATCAGTTTGTATGGCCAAATGCGGTATCAACAACCTGCAACGACGAT 708
DB 2790 CGGCGCGAGAACCGCGCGCCGATTAAGATCAGTGAATGATGACGAGTCTGCTCAAGAT 2731

QY 709 GATCAGCGGCACACTTTGGGCGTGGACAGCGGGGCTGACATG 752
DB 2730 AGCCCATCAAAATCAGAGAGCACAAAGCTCAGATGGAATG 2687

```

RESULT 8

```

US-08-362-577C-20/c
; Sequence 20, Application US/08362577C
; Patent No. 5807673
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eba, Soji

```

```

; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,577C
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: Clinical Isolate EC-625
; US-08-362-577C-20

Query Match 4.2%; Score 100.8; DB 1; Length 5541;
Best Local Similarity 50.5%; Pred. No. 2.2e-19;
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

QY 1 AGATATCTCTTTGGAAGAACCATGACGATTCGCTGACATTTGCGTCTGGAAGGC 60
DB 3416 AATATACCGCATGGAAGAACCGCTCTGCGTGGCTCATGCGGTACAAAGCGGTAAGC 3357

QY 61 TCCTTACGTGGTATTTCTTCTACGGTCCAGAGCTCAGCGAGGCGGTGATGAT 120
DB 3356 GCTTTATGTGCGGATCTCTCTTACTCGCAGAGCGGCAAAAATGTTGATGTCT 3297

QY 121 GCGGAGAGGCGTCCCGCTTCTGATTCTATCAGCAAGCTATTCATCATTAATCGTTG 180
DB 3296 GCGGAGTGGAAAATTCGCTGTTTAACTTCACTTCTGTAACAATTTACTGAACCGCTG 3237

QY 181 GGTGAGAACCGCGGATGACGTGAGAACCTTGTGACGTGAGTCCCAATGCTCT 240
DB 3236 GGTGAGTAAAGCGG-----CCTGCGATACCTTCGCAAAATAACGCGCT 3192

QY 241 TGGGCTCATTTCTTCTCACCCTTGGCAGGCGCTGCTCAGCGACAAATATCTCGATGG 300
DB 3191 GGGCTGATTTGCTTCTTCTCTGCGTCAAGGATTTGCTGACCGAATAATCTCAACGG 3132

QY 301 AATTCAGAGGCTTCCCG-----CGCCAGCGAGGTAAGTCCCTGTCTGAGGG 348
DB 3131 CATTCGCAAGATTCACGATGATCGTGAAGGAATAAAGTTCTGTTGCTGACCGGAA 3072

QY 349 CATGTTGAACGTGAACATATTTGATGTTGTCGCAAGCTCAATGACATCGCCAGGAACG 408
DB 3071 AATGCTTACCGAAGCAACCTCAACAGCCTACATTTATTGAATGAATGGCACAGACGG 3012

QY 409 CGGCGAGTCACTTTGGCGAGATGGCGCTTTCATGGGTGCTGCGGAGCAAGGAGTACGG 468
DB 3011 TGGCAATCAATGGCGCAATGGCGTTAAGCTGTTGCTGAAAGA-----2967

```

QY 469 CGCGGATACCGTGACAGTGCATTTGATTTGTTCTGTCAGTTGACGAGTGCACACAG 528
Db 2966 ---TGATCGCTGACGTCGGTATTGATTTGTC- CAGCGCGGAGCAACTTGAGAGAA 2911
QY 529 CTTGATTACTCAACACTTGGAGTTTCTGACCGCGAGTTGGAGCGGATCGATGAGAT 588
Db 2910 CGTGACGCGCTGAATAATCTGACATTTAGCACCGAGGAGCTGGCGCAGATTGATCAGCA 2851
QY 589 TTCCACGACCGCGGATCAACATTTGGGCGAGGCCACCGATTCCAAACCGCGGAAA 648
Db 2850 TATCCGCGATGGGAGCTGAATCTGTGCGAGGCGTCTTCGGATAAATGACCTGTGTAATA 2791
QY 649 CTAACCCATCAACATCAGTTTGTATGGCCATGCGGTCATCAACACTGCCACGACGACTT 708
Db 2790 CGGCGCGGAGAACCGCGCGGATAGATCAGTATGACGAGTCACTCGGTCAAGAT 2731
QY 709 GATCCAGCGCACACTTGGGCGTGGACAGCGGCGGTGACAAATG 752
Db 2730 AGCCCATCAAAATGCAGAGAGACAAACGTCAGATGGATAATG 2687

RESULT 9

US-08-920-828-20/c
; Sequence 20, Application US/08920828
; Patent No. 5853998
; GENERAL INFORMATION:
; APPLICANT: Obno, Tsuneya
; APPLICANT: Matshuisa, Aki
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,828
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: Clinical Isolate EC-625
US-08-920-828-20

Query Match

4.2%; Score 100.8; DB 2; Length 5541;

Best Local Similarity 50.5%; Pred. No. 2.2e-19;
Matches 386; Conservative 0; Mismatches 332; Indels 46; Gaps 4;

QY 1 AGATACTCCTTTTGAAGAAGAACCATGTACGATTCGCTGATGATTTGGTCTGGAAGGC 60
Db 3416 AATATCGCGGATGGAAGAAGAACCGCTCTCGCTGCTCATCGGTACAAAGCGGTGAGGC 3357
QY 61 TCTTTACGTGGGTATTTCTTCTTACGCTCAGAGCTCAGCGGAGCGGCTGAGTTCAAT 120
Db 3356 GCCTTTATGTCGGGATCTCTCTTACTTCGCGCAGAGCGGCAAAAATGTTGAGTTGCT 3297
QY 121 GCGGAGGAGGAGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCATTAATCTGTTG 180
Db 3296 GCGGAGGAGGAGGATTCGCTGTTAATTCATCAACCTTCGTACATTTTACTGAAACGCTG 3237
QY 181 GGTGGAGGAGGAGGCGGATGAGCGTGAGAACTTTGTGAGTGTGCTGCAACAAATGCTCT 240
Db 3236 GGTGGAGGAGGAGGCGGATGAGCGTGAGAACTTTGTGAGTGTGCTGCAACAAATGCTCT 3192
QY 241 TGGCGTCAATTCCTTTCTACCACTTGGCGAGGCGCTGCTCAGGACAAATATCTCGATGG 300
Db 3191 GGGCTGATTCCTTTTACTCTCTGCTCAGGAGTGTGACCGGAAATATCTCAACGG 3132
QY 301 AATTCCAGAGGTTCCCG-----CGCCAGCGAGGTTAAGTCCCTGTCTGAGGG 348
Db 3131 CATTCGCGAAGATTTCAGGATGCTGTAAGGGAATAAAGTTCTGTGCTGAGCGCGAA 3072
QY 349 CATGTTGAAGAGTGAACAAATATGATATGTCGCGAGCTCAATGATCATCGCCGAGGAGC 408
Db 3071 AATGCTTACCGAAGCCAACTCAACAGCTTACACTTATTGAATGAATGGCAGCAGCG 3012
QY 409 CGGCGAGTCACCTGCGCAGATGCGCTTGCATGGGTGCTGCGCGAGCAAGGAGTACGG 468
Db 3011 TGAACATCAATGCGCAATGCGGTTAAGCTGTTGCTGAAAG----- 2967
QY 469 CGCGGATACCGTGACAGTGCATTTGATTTGCTGCTGCTGAGTGTGAGCAGCTGGACAAAG 528
Db 2966 ---TGATCGCTGACGTCGGTATTGATTTGCTGTC- CAGCGCGCGGACAACTTGAGGAGAA 2911
QY 529 CCTTGATTCATCAACACTTGGAGTTTCTGACCGCGAGTTGGAGCGGATCGATGAGAT 588
Db 2910 CGTGACGCGCTGAATAATCTGACATTTAGCACCGAGGAGCTGGCGCAGATTGATCAGCA 2851
QY 589 TTCCACGACCGCGGATCAACATTTGGGCGAAGCCACCGATTCCAAACCGCGGAAA 648
Db 2850 TATCGCGGATGCGGAGCTGATCTGTGCGAGCGCTTTCGGATAAATGACCTGTGTAATA 2791
QY 649 CTAACCCATCAACATCAGTTTGTATGGCCATGCGGCTGATCAACACTGCCACGACGAGCTT 708
Db 2790 CGGCGCGGAGAACCGCGCGGATAGATCAGTATGACGAGTCACTCGGTCAAGAT 2731
QY 709 GATCCAGCGCACACTTGGGCGTGGACAGCGGCGGCTGACAAATG 752
Db 2730 AGCCCATCAAAATGCAGAGAGACAAACGTCAGATGGATAATG 2687

RESULT 10

US-09-489-039A-5370
; Sequence 5370, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5370
; LENGTH: 1095
; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5370

Query Match	3.4%;	Score 81.4;	DB 4;	Length 1095;
Best Local Similarity	51.5%;	Pred. No. 5.2e-14;		

QY	1	AGATCTCTCTTTGGAGAAACCATGTAGCATTCGCTGACCATTTGTTGCTTCGGAAGGC	60
Db	528	ACAGACGCCCTCATGGAAACGATGCGTGCCTGCATCATCTGGTTCGTCAAGGTTAAAGC	587
QY	61	TCTTTTACGTGGGTATTCTTCTTACCGTCCAGAGCTCAAGCGGAGCGGCTGAGTTCAT	120
Db	588	GCTATATGTCGGGATCTCCAACTATCTCTGSCCCAGGCGCGGAGCGGCTGAAGATCCT	647
QY	121	GCGGAGGAGGGCTGCCGCTTCTGATTCATCAGCCAAAGCTATTCGATCATTAATTCGTG	180
Db	648	TAAAGCACTTCGGCACCCCTGTCATATTACACGCTTCGCTACTCGATGTTTCGAAACGGG	707
QY	181	GGTGAGAGAAACCGGCGGATGACGCTGAGAAATCTTTGTGAGTCAGCTCCCAACAAATGGTCT	240
Db	708	CTGTGGAAGAAGG-----TTCTGGATTTTCTGCAGACGGAAGGAT	749
QY	241	TGGGCTATGTCCTTCTCACATCTGCGCAGGGCTGTCTCAGGACAAATATCTCGATGG	300
Db	750	TGGAGCATAGCGTTTTCACGCTGCGCGGCGGCAACTTACCAGCGCTATCTCAACGG	809
QY	301	AAATCCAGAGGGTTCCCG---CGCCAGCCAGGGTAAGTCCCTGTCTGAGGGCATGTTGAA	357
Db	810	CAATCCGGCGCACTCGCGCGCGCAGCAGCATGCTGTTTTCTGCAGCCGAGCACTGAC	869
QY	358	CGTGAACAAATTTGATATATGTGTCGCAAGCTCAATGACATGCCCAAGAGACGGCGGAGTC	417
Db	870	TCCAGCCCGATTAGAAAAAATTTGTCAGCTTAAACGSCAGGCGGAAGCGCTGGGCGAA	929
QY	418	ACTTCGCGCAGATCGGCGCTTGCATGGGTGTCGGCGGACGAGGAGTACGGCGCGGATAC	477
Db	930	GCTGTGCGAGATGGCGCTGGGCTGGGTGTTGGCGAAGAGAAAGTCACTTCGGTCTCAT	989
QY	478	CGTGACCAAG	486
Db	990	CGGCGCCAG	998

RESULT 11
 US 09-489-039A-4735
 ; Ser. No. 4735 Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary B'reton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

Query Match	3.4%;	Score 80.2;	DB 4;	Length 915;
Best Local Similarity	44.8%;	Pred. No. 1.1e-13;		
Matches 396;	Conservative 0;	Mismatches 478;	Indels 9;	Gaps 2;

QY	1423	GAAACCCCAATCAACTCGGACATTTTGCTCTCAATCATTTGATGAAGCAGCTTCGAAGCGC	1482
Db	33	GGACTACAGGACACTACAAGCGTGGATCGTGATTAGGACACGAGGATTTGACGGCGC	92

QY	1483	CTCCTTAGCCCTTTCCATTCCCTCGGCGGTGAGTCAGCGCGTTAAAGCTCTCGAGCA	1542
----	------	---	------

Db	93	CGCGCAGAGCTATGATTATACCCAGTCCCGCGTCTCAGAGCGTATTAGCAGCTGGAATA	152
Qy	1543	TCACTGGGTCGAGTGTGGTATCGCGCACCCAAACGGGCCAAAGCAACCGAAGCGGGTGA	1602
Db	153	TATGTTGGCCACGCCGCTGCTGGTGACGTGCGCGCGCCCAACCGAAGCA	212
Qy	1603	AGTCCTTGTGCAAGCAGCGGGAATAATGGTGTGCTGACAGCAGAAACTTAACGGCACT	1662
Db	213	GAATCTGCTGGCTCTGTGGCCAGGTTGAATCTCTGGAGGAAGTGGCTGGGGCGATGA	272
Qy	1663	ATCTGGAGCCCTTGTGAAATCCCGTTTACCATTGCCATCAACGCGAGATTTCGTATCCAC	1722
Db	273	ACAAACCGGCTCTAGCGCGTTGCTGTGCTGGCGGTGAACGCCGACAGCTCTGCGAC	332
Qy	1723	ATGGTTTCTCCCGTGTTCAACAGAGTAGTCTTCTTGGGGTGGAGCAAGCTCACGCTCGG	1782
Db	333	CTGGCTGTGCGCGCGCTGGCCAAAGTTCGTGACATCCCTTATTCGTCTCAACTGCA	392
Qy	1783	CTTGAAGATGAAGGCGCACATTAATTCCTTGTGCGCGGTGGAGATGTTTTAGGAGCGGT	1842
Db	393	GGTGAAGATGAATCCCGCACACGAGAGCGCTCGCGCGGTGGCGAAGTAGTGGGCGCAGT	452
Qy	1843	AACCCGTGAAGCTAATCCCGTGGCGGATGTGAAGTAGTAGAATTTGAAACCATGGCCCA	1902
Db	453	GAGTATCCAGCGCAGCGCGCTGCCAAGCTGCTGCGTGCATCAGCTTGGGGGCGCTGATTA	512
Qy	1903	CTTGGCCATTGCAACCCCTCATTTGCGGAGTGCCTACATGTTGATGGAAACTAGATTG	1962
Db	513	CTGTGTTGTCGATCAAGAAGAGTTTGGCGAAGCTATTTCCCGATGGCGTGACGGTTC	572
Qy	1963	GGCTCGCAT-----GCCCGCTTTAAGCTTCGGTCCCAAGATGTGCTTCAAGACCGTGA	2016
Db	573	GGCGCTGCTTAAGGCGCGGTGCTGCGCTTTGACCATCTGACGATATGCATCAGGCGTT	632
Qy	2017	CTTGAACGGGCGGTGCATATGTTGTTGGGCGCAGCGGGGTATCCATGTGTCGCTCGCG	2076
Db	633	CTTGCAGCAAAATTTCGACTCTGCGCGCGGCGAGCGTCCCTGCGCACATCTGCTCAACTCGTC	692
Qy	2077	GGAGGTTTTGTGAGGCATTTGCGCGAGGCTTGGTTGGGACCTCTTCCCGAAACCCA	2136
Db	693	GGAGCTTTGTGCACTGTGGCGCGCAGGGCACCACTGCTGTATGATCCCGCATCTCGA	752
Qy	2137	AGCTGCTCCCATGCTAAAGCAGGAGAGTGTATCTCTCGATGAGATAGCCATTGACAC	2196
Db	753	GATCGAAAAGAGCTCAACAGGCGAATTTATCGATCTCACCCGGGCGCTGTTCCAGCG	812
Qy	2197	ACCGATG---TATTTGGCAACGATGGCGCTTGGAACTTAGATCTTAGCTAGATTCAAGA	2253
Db	813	CGCATGCTCTACTGGCACCGCTTGCGCCCGGAAAGCGCATGATGCGCGGGTGAACGA	872
Qy	2254	CGCGCTCGTTGATGACGCAATCAGAGGATTCGGGCTTAGTTA	2296
Db	873	TGCGGCTCAATGACTACGACACAAGTGTGCTGCATCAGGATTA	915

RESULT 12

US-09-543-681A-791
; Sequence 791, Application US/09543681A

; Patent No. 6605709
; GENERAL INFORMATION:

APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAE
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128.706

;
;
;

ORGANISM: Proteus mirabilis
US-09-543-681A-791

Query Match 3.1%; Score 73.4; DB 4; Length 936;
Best Local Similarity 44.9%; Pred. No. 1.2e-11;
Matches 324; Conservative 0; Mismatches 391; Indels 6; Gaps 1;
QY 1455 TCATTGATGAGGAGGCTCGAAGCGGCTCTTAGCCCTTTCATTTCCCTCGCGG 1514
Db 80 TTATTCGAGAGCGTGGTTTGAACGAGCGGCAAAACTATGATTACCAATCTGCG 139
QY 1515 TGAGTCAGCGGTTAAAGCTCTCGAGCATCACTGGGTCGAGTGTGATCGCCACC 1574
Db 140 TTTCACAAAGAAATAAACCACTTGAGATCTATTTCGACAACTTTATTAGTAGCTAG 199
QY 1575 AACCGCCCAAGCAAGCGGTCGAGTCTCTGTCGAGCGCGGCAAAATGGTGT 1634
Db 200 TTCGCGCCCAACCCACAGAGCGAGGGCAAAAGCTATCGCATTAAGTCGAA 259
QY 1635 TGTCGCAAGCAAGAACTAAAGCGCACTATCTGAGCGCTTGTGAAATCCGTTAACCA 1694
Db 260 TGTAGAGAGCACTGTTAGGTGATGAAATAGTGTCTGCGCGCTTTTACTCTTT 319
QY 1695 TCGCATCAACGAGATTCGCTATCCATGTTTCTCCGCTGTTCAACGAGTAGTGT 1754
Db 320 TAGCTGTCAACGCGGATGTTTAGCTACTTGGTTATTGCTGCTTTCACCGGTATTA 379
QY 1755 CTTGGGTTGAGCAAGCTCAGCTGCGCTTGAAGATGAGCGCACACATTTCTTGC 1814
Db 380 CACACTTCTATCCGCTTAATAATCAAGTAGAATGAACTCGTACCAAGAGCGAT 439
QY 1815 TCGCGCTGAGAGATGTTTGGAGCGGTAAACCGTGAAGCTAATCCGCTGGCGGATGT 1874
Db 440 TAAGACGTTGAGTGGTGGTGTATCAGTATTCAGCCCCAAGCAATGCTAGCTGTC 499
QY 1875 AAGTAGTAGAATTTGAACATGCGGCACCTTGGCCATTTGCAACCCCTCATTTGCGGATG 1934
Db 500 TAGTCGACCAATAGGCGAATPAGACTATCTTTTGTGGCATCCCTGATTTTGCCCAAC 559
QY 1935 CTTACATGTTTGTATGG3-----AACTAGATTGGCTGGCATCCGCTTACGCTTCG 1988
Db 560 GCTATTTTGCMAATGGGGTCACTAAATCATCTCTGTTAAAGCCCGCTGCGCATTTG 619
QY 1989 GTCCAAAGATGTCTTCAAGACCGTGACCTGAGCGGCGGTGATGCTCTGTGGGCG 2048
Db 620 ACCATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 679
QY 2049 GCAGCGCGTATCGATGTCGCGCGGGAAGGTTTGGTAGGCAATTCGCGGAGGCC 2108
Db 680 GTGTGCGCGTGCATATTTGTTAATCTTCCGAGCGCTTTGTTCAAGTAAACAAGGCT 739
QY 2109 TTGGTTGGGAGCTTCTTCCGAAACCCAGCTGCTCCCATCTAAAGCAGGAGATGA 2168
Db 740 CACATGTTGTATGATCCCTCATCTACAAATTCGATGAGCTAAAGTGGTGAATGA 799
QY 2169 T 2169
Db 800 T 800

RESULT 13
US-09-252-991A-7644
Sequence 7644, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7644
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7644
Query Match 2.7%; Score 65.2; DB 4; Length 900;
Best Local Similarity 46.2%; Pred. No. 3.2e-09;
Matches 415; Conservative 0; Mismatches 458; Indels 25; Gaps 5;
QY 1418 TTCATGAACCCCATTAACATGAGACATTTCTCTCATCATTTGATGAGGCGAGTCGAA 1477
Db 3 TTGTTGCACTACAGTTGCTCGCGCGCTGCGCGGATTCGACGCGATCAGCTGCTC 62
QY 1478 GGGCGCTCTTAGCCCTTTTCCATTTCCCTCGCGGTCGAGTCAGCGGTTAAAGCTCTC 1537
Db 63 CGCGCGCACAGCGCTGGGGCTGTGCAATCGCGGATTCGACGCGATCAGCTGCTC 122
QY 1538 GAGCATCAGTGGGTGAGTGTGATTCGCGCAACCCAAACCGGCGCAAGCAACCGAAGCG 1597
Db 123 GAGCGCGCGGTGCGCGAGCGGCTCTGTTGCGCGAGACCGCGCATCCCAACCGATCTC 182
QY 1598 GGTGAAGTCTTGTGCAAGCAGCGGAAATGTTGTTGTCGACGAGCAAACTAAAG-- 1655
Db 183 GGGCGCGCTGTTTCAACCATGTCAGAGTGGGCTGCTGAGGGGAGCTGCGAGCGC 242
QY 1656 -----CGCAACTATCTGAGCGCTTCTGAAATCCCGTTAAACCATCGC-CATCAACGC 1707
Db 243 TGGGTGCGCAACTGATGAAGCGGTGCCCCGAGCGCTCGCATCGTCTCAACGC 302
QY 1708 AGATTGCTATTCACATGTTTCTCCGCTTTCACGAGGTAGCTTCTTGGGGTGGAGC 1767
Db 303 CGATAGCTTGGCACTTGGTGGCGCGCGGTGGCGGATTTCTGCGCGAGCGGGGT 362
QY 1768 AAGCTTCACGCTGCGCTTGAAGATGAAGCGCACACATATCTTCTGCTGGCGGTGAGCA 1827
Db 363 GCTGCTGAGCACTTGTGTGAGGACAGAGGTGGGCTCAAGCGCATGCGCGGGGGA 422
QY 1828 TGTGTTAGAGCGGTAAACCGTGAAGCTTAATCCGCTGGCGGAGATGTAAGTAGTAGAAT 1887
Db 423 AGTGGCGGCTGCTCTGCGGTAGCGCGCGCGGTGGCGCGGCTTCTGCTGCTGCT 482
QY 1888 TGGAACCATGCGCACTTGGCCATTTGCAACCCCTCATTT-----GGGGATGCTACAT 1941
Db 483 CGAGCCATGCTGTTATCGCGGCTTGGCCAGCGCGGATTTTCAATGCGCGGACCTTTCCCG 542
QY 1942 GGTTCATGGGAACTAGATTTGGGCTGCGATGCGCGCTTTACCGCTTCCGTCCTCAAGATGT 2001
Db 543 CGGGTTCGAGCGCGCGCTGCGCGGCTGCGCGGATCGTGTGTCGCGCGGAGCGCT 602
QY 2002 GCTTCAAGACGCTGACCTGAGACGGCGCGCTGATGCTCTGTTGGGCGCGAGGCGGATTC 2061
Db 603 GCTTCAACACCGCTTTCTCAAGGACCTCGCG-----TCGAGGCGGCTTTTCAATCCACCA 656
QY 2062 CATTTGTCGCGCGGAGGTTTGGTGGGCAATTCGCGAGGCTTGGTGGGAGCT 2121
Db 657 CTTTGGCGCGCTTGGGAGGTTTCTGCGCGCTTCAACCGCGCGGCTTCTCGGCTGGGCGCT 716
QY 2122 TCTTCCGAAACCCCAAGCTGCTCCCATGCTAAAGCAGGAGAGTGTATC---CTCTCGA 2178
Db 717 GGTTCGCGAGCGCGAGTGAAGCGGAGCTGGCGCGCGGTGAATCTGTGAGCTGTGCTCC 776
QY 2179 TGAGATACCCATTGACACCGATGTTATGGCAACGATGGCGCTTGAATCTTAGATCTCT 2238
Db 777 CGGCCAGGTCTATGACGATGATCCGTTGTTACTGCGCACTACTGCGCAACCGCGGCGAATGCT 836
QY 2239 AGCTAGACTCACAGCGCGCTGTTGATGAGCAATTCGAGGAGTTCGGGCTTAGTTA 2296
Db 837 CGGCTGCTACCGAGACCTCTCTGCGCGCGCGGAGCGCTTGGTGGGCTGTCA 894

RESULT 14

US-09-724-623-24
 ; Sequence 24, Application US/09724623
 ; Patent No. 6476209
 ; GENERAL INFORMATION:
 ; APPLICANT: Glenn, Matthew
 ; APPLICANT: Lubbers, Mark W
 ; APPLICANT: Dekker, James
 ; TITLE OF INVENTION: Polynucleotides, materials incorporating
 ; TITLE OF INVENTION: them, and methods for using them.
 ; FILE REFERENCE: 1048U1
 ; CURRENT APPLICATION NUMBER: US/09/724,623
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 124
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24
 ; LENGTH: 1002
 ; TYPE: DNA
 ; ORGANISM: Lactobacillus rhamnosus
 US-09-724-623-24

Query Match 2.7%; Score 64; DB 4; Length 1002;
 Best Local Similarity 47.9%; Pred. No. 7.8e-09;
 Matches 279; Conservative 0; Mismatches 270; Indels 33; Gaps 2;
 QY 9 CTTTGAAGAAACCATGTACGCAATTGCGTGTGATTTGCGTCTGGAAGGCTCTTTACG 68
 DB 446 CTTTGAAGAGACGTTAATGCGTGTGATCAACGGTGTGACGTGAAGCGTTGTATA 505
 QY 69 TGGGTATTTCTTCTACGGTCCAGAGCTCACAGCGGAGGCGGTGAGTTTCATGCGGAG 128
 DB 506 TTGGTATTTCCAACTATGATAGCAAGCAGACCAAGAGCAATTCGGATGTTTAAAGATC 565
 QY 129 AGGCTGCGCGCTTCTGATTCATCAGCCAGCTATTCATCATTAATCGTTGGGTGAGG 188
 DB 566 TGCACACGCGCTTTGTGTAATCAATCAATCAGTTAACAATGTTTAAATCGACCGCTGAA 625
 QY 189 AACCGGCGATGACGGTGAGAACTTCTGTCAGTCAGTGCACCAATCGTCTTGGCGTCA 248
 DB 626 GTCGCGG-----CTTGATCGATGATTAAGAGCTGATGTCGCGGTTGA 570
 QY 249 TTGCTTTCTCACCATTTGCGGAGGCGCTGCTCAGCGACAAATATCTGATGGAATTCAG 308
 DB 671 TTGCATACGACCGGTTATCAGAGGCTTGTATCAGATCGCTACCTAAAGGGAATTCGCG 730
 QY 309 AGGGTTCGCGCGCCAGCGGTAAGTCCCTGCTGAGGCGATGTTGAACGTGAACATA 368
 DB 731 ATACTTTCAAAATCCATCCAAACAGGCCACTTTTGAAGGCAAGAGGCTGTGG 790
 QY 369 TTGATATGCTCCGCAAGCTCAATGACATCGCCAGGAACCGGGCAGTCACTTGGCGAGA 428
 DB 791 TTAAGCAACTAAATCGGCTTAATGAATTTGCGCATGATCTGACCAACCTGAGTCAAA 850
 QY 429 TGGCGCTTGATGGGTGCTGCGGAGCAAGAGAGTACGGCGCGGATACCGTGACCACTG 488
 DB 851 TGGCGCTTGGCGGTGTTTACGAGATCCGG-----TTGTCACAAGTG 892
 QY 489 CATTTGATGGTCTTCTGTCAGTTGAGCAGCTGGACACAGCCTTGATTCACCTCAACA 548
 DB 893 TGATCATTTGGAGACACCTCAGTTGAACACCTTCAGGATAACCTTAAAGCAACGACATC 952
 QY 549 TGGAGTTTTCTGACCGGAGTTGGAGCGCATGATGAGATTT 590
 DB 953 TGACCTTTTACTGCTGAAGAGATTCAACAAATTGATGATATTT 994

RESULT 15

US-09-252-991A-7908/c
 ; Sequence 7908, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 7908
 ; LENGTH: 699
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-7908

Query Match 2.6%; Score 61.8; DB 4; Length 699;
 Best Local Similarity 54.6%; Pred. No. 2.9e-08;
 Matches 148; Conservative 0; Mismatches 117; Indels 6; Gaps 1;
 QY 650 TAACCATCAACATCAGTTTGTATGCGCAATGCGGTCTCATCACTGCCACGACGACGTTG 709
 DB 698 TATCCCGGAGACAGCAGTTGCGCGGCCATGCCAGCATCATGGCCGCCACCATCAGTCCG 639
 QY 710 ATCCAGCGCCACACCTTTGGGGCTGGACAGCGGGCGTGACATCTCTGCTGCCGAAACCC 769
 DB 638 AGCAGGCGCCAGGTGGCCCGCGCGCCAGCCAGAGGGGCGACCATGCCGCGCGAGGGCG 579
 QY 770 ACCAGCGGGAACAGATCAGGCTTGGCGGACGCGGCGGCAAAATCCACGCTCCG 829
 DB 578 AGGCGGAAGAACCAATCAGCAGGCGCTTGGCGCGCGCGGAGGCGATAGGGCGCGCGCG 519
 QY 830 GTGTGCGCGGTATTGCGCGCGAGCGCGCGGATTAACACAAACCGCTCCAAATACGCAATC 889
 DB 518 G-----CCTGCTGGCGCGCGGAGCGACCGATCAGAGTAGCGTATCGAGATAGACGTTG 465
 QY 890 GGGTTCAACAGGTGACGACGATTTGCCATCA 920
 DB 464 GGGTTGAGCAGGCGTACCGCCAGTGGCGCCA 434

Search completed: March 16, 2004, 03:01:32
 Job time : 197.837 secs

Result No.	Score	Query Match	Length	DB	ID	Description
C	1	2374	100.0	2374	14	US-10-196-232-24
	2	2374	100.0	3309400	9	US-09-738-626-1
	3	993	41.8	993	10	US-09-746-660A-47
C	4	870	36.6	870	9	US-09-738-626-3456
	5	822	34.6	822	10	US-09-746-660A-51
	6	711	29.9	711	14	US-10-166-142-9
C	7	708	29.8	708	9	US-09-738-626-3455
	8	696.8	29.4	712	14	US-10-166-142-9
	9	627	26.4	627	9	US-09-738-626-3454
C	10	485.2	20.4	993	12	US-10-627-476-669
	11	305.8	12.9	1095	12	US-10-282-122A-17521
	12	302.8	12.8	879	12	US-10-282-122A-17523
C	13	164.4	6.9	1026	15	US-10-369-493-3537A
	14	164.4	6.9	1026	15	US-10-369-493-38826
	15	164.4	6.9	1028	15	US-10-369-493-38124
C	16	164.4	6.9	1026	15	US-10-369-493-38124
	17	164.4	6.9	1028	15	US-10-369-493-38124
	18	164.4	6.9	1026	15	US-10-369-493-38124
C	19	164.4	6.9	1028	15	US-10-369-493-38124
	20	164.4	6.9	1026	15	US-10-369-493-38124
	21	164.4	6.9	1028	15	US-10-369-493-38124
C	22	164.4	6.9	1026	15	US-10-369-493-38124
	23	164.4	6.9	1028	15	US-10-369-493-38124
	24	164.4	6.9	1026	15	US-10-369-493-38124
C	25	164.4	6.9	1028	15	US-10-369-493-38124
	26	164.4	6.9	1026	15	US-10-369-493-38124
	27	164.4	6.9	1028	15	US-10-369-493-38124
C	28	164.4	6.9	1026	15	US-10-369-493-38124
	29	164.4	6.9	1028	15	US-10-369-493-38124
	30	164.4	6.9	1026	15	US-10-369-493-38124
C	31	164.4	6.9	1028	15	US-10-369-493-38124
	32	164.4	6.9	1026	15	US-10-369-493-38124
	33	164.4	6.9	1028	15	US-10-369-493-38124
C	34	164.4	6.9	1026	15	US-10-369-493-38124
	35	164.4	6.9	1028	15	US-10-369-493-38124
	36	164.4	6.9	1026	15	US-10-369-493-38124
C	37	164.4	6.9	1028	15	US-10-369-493-38124
	38	164.4	6.9	1026	15	US-10-369-493-38124
	39	164.4	6.9	1028	15	US-10-369-493-38124
C	40	164.4	6.9	1026	15	US-10-369-493-38124
	41	164.4	6.9	1028	15	US-10-369-493-38124
	42	164.4	6.9	1026	15	US-10-369-493-38124
C	43	164.4	6.9	1028	15	US-10-369-493-38124
	44	164.4	6.9	1026	15	US-10-369-493-38124
	45	164.4	6.9	1028	15	US-10-369-493-38124
C	46	164.4	6.9	1026	15	US-10-369-493-38124
	47	164.4	6.9	1028	15	US-10-369-493-38124
	48	164.4	6.9	1026	15	US-10-369-493-38124
C	49	164.4	6.9	1028	15	US-10-369-493-38124
	50	164.4	6.9	1026	15	US-10-369-493-38124
	51	164.4	6.9	1028	15	US-10-369-493-38124
C	52	164.4	6.9	1026	15	US-10-369-493-381

Qy	1	AGATATCTCTTTGGAGAAACCAATGATCCCATTCGGTGACATTTGTCGCTCGAAAGCG	60
Db	2374	AGATATCTCTTTGGAGAAACCAATGATCCCATTCGGTGACATTTGTCGCTCGAAAGCG	2315
Qy	61	TCCTTTAGTGGGTATTTCTTCCTACGGTCCAGAGCTCACACGGAGGGCGGCTGAGTTTCAT	120
Db	2314	TCCTTTAGTGGGTATTTCTTCCTACGGTCCAGAGCTCACACGGAGGGCGGCTGAGTTTCAT	2255

QY 121 GCGGAGGAGGCGCGCTCTGATTCATCAGCAGAGCTATCCATCATTAATCGTTG 180
Db 2254 GCGGAGGAGGCGCGCTCTGATTCATCAGCAGAGCTATTCATCATTAATCGTTG 2195
QY 181 GGTGGAGGAACCGGCGATGACGGTGAGAACTTGTTCAGTCAGCTGCGCAACATGGTCT 240
Db 2194 GGTGGAGGAACCGGCGATGACGGTGAGAACTTGTTCAGTCAGCTGCGCAACATGGTCT 2135
QY 241 TGGGTCATGCTTCTTCACACCTTCGCGCAGAGGCGCTCTCAAGGACAAATATCTCGATGG 300
Db 2134 TGGGTCATGCTTCTTCACACCTTCGCGCAGAGGCGCTCTCAAGGACAAATATCTCGATGG 2075
QY 301 AATTCCAGAGGGTTCGCGCGCAGCGAGGTAAGTCCCTGTCTGAGGCGCATGTTGAACGT 360
Db 2074 AATTCCAGAGGGTTCGCGCGCAGCGAGGTAAGTCCCTGTCTGAGGCGCATGTTGAACGT 2015
QY 361 GAACAATATTTGATATGTTGTCGCAAGCTCAATGACATCGCCAGGAACCGCGGCAATCACT 420
Db 2014 GAACAATATTTGATATGTTGTCGCAAGCTCAATGACATCGCCAGGAACCGCGGCAATCACT 1955
QY 421 TGGCAGATGGGCGCTTGCATGGGTGCTGCGGAGCAGAGAGTAGCGCGGATACCTT 480
Db 1954 TGGCAGATGGGCGCTTGCATGGGTGCTGCGGAGCAGAGAGTAGCGCGGATACCTT 1895
QY 481 GACCAATATTTGATATGTTGTCGCAAGCTCAATGACATCGCCAGGAACCGCGGCAATCACT 540
Db 1894 GACCAATATTTGATATGTTGTCGCAAGCTCAATGACATCGCCAGGAACCGCGGCAATCACT 1835
QY 541 CAACAATTTGAGATTTGTCGCGCGAGTTGAGGCGCATCGATGAGATTTCCACGAGC 600
Db 1834 CAACAATTTGAGATTTGTCGCGCGAGTTGAGGCGCATCGATGAGATTTCCACGAGC 1775
QY 601 CGGCATCAACATTTGGGAGGAGGCGACCGATTTCCAAACCCGGAACCTAACCATCAA 660
Db 1774 CGGCATCAACATTTGGGAGGAGGCGACCGATTTCCAAACCCGGAACCTAACCATCAA 1715
QY 661 CATCAGTTTGTGCGCAATGGGTTCATCAACCTGCCACGAGAGTTGATCCAGCGCCA 720
Db 1714 CATCAGTTTGTGCGCAATGGGTTCATCAACCTGCCACGAGAGTTGATCCAGCGCCA 1655
QY 721 CACTTGGGCGTGGACAGCGGCGTGCACAAATGCTGTCGCGCGAAACCCACGAGCGGAA 780
Db 1654 CACTTGGGCGTGGACAGCGGCGTGCACAAATGCTGTCGCGCGAAACCCACGAGCGGAA 1595
QY 781 CCAGATCAGGCTTTCGCGGAAACCGCGAGCGGAAATCCACGCTCGGCTGCGCGTA 840
Db 1594 CCAGATCAGGCTTTCGCGGAAACCGCGAGCGGAAATCCACGCTCGGCTGCGCGTA 1535
QY 841 TTGCGCGCGACGCGCGGATTAACAAACAAACGCGTCCAAATACGCAATTCGGGTTCAACCA 900
Db 1534 TTGCGCGCGACGCGCGGATTAACAAACAAACGCGTCCAAATACGCAATTCGGGTTCAACCA 1475
QY 901 GGTTCAGCAGATTTGCGCATCAACATGGGCTTTTACCAACCGGCTGCTATCGAGCTCAC 960
Db 1474 GGTTCAGCAGATTTGCGCATCAACATGGGCTTTTACCAACCGGCTGCTATCGAGCTCAC 1415
QY 961 CTCACACCGCACCGGTTGCGGTCAGTGCGCACCGCGCAACCGCGGCGGCTGTC 1020
Db 1414 CTCACACCGCACCGGTTGCGGTCAGTGCGCACCGCGCAACCGCGGCGGCTGTC 1355
QY 1021 ATCGGCGACGGTGTGTTCTTCAATGATCTGTGGGCTTCCACCTTGTGTCAT 1080
Db 1354 ATCGGCGACGGTGTGTTCTTCAATGATCTGTGGGCTTCCACCTTGTGTCAT 1295
QY 1081 GGCCTCTTCTGCGCATGACGCGCAACCATTAACAGTAAAGTAAAGCGGCTGTC 1140
Db 1294 GGCCTCTTCTGCGCATGACGCGCAACCATTAACAGTAAAGTAAAGCGGCTGTC 1235
QY 1141 AATATCAGACGATCGGCGGCGATTTGGACAAAGATTCACACCGCCACAGGTGCGCGGAT 1200
Db 1234 AATATCAGACGATCGGCGGCGATTTGGACAAAGATTCACACCGCCACAGGTGCGCGGAT 1175
QY 1201 GAACAAAAGAGCGTCAGAAATTAACACACGAGAGAACCGCAATGAGTCTTCGCGCTT 1260
Db 1174 GAACAAAAGAGCGTCAGAAATTAACACACGAGAGAACCGCAATGAGTCTTCGCGCTT 1115
QY 1261 AATTCTTGTGTTAATCACAGTACATTCCTGCGTCCGATCGACAGTAAAGACTGGCCCC 1320
Db 1114 AATTCTTGTGTTAATCACAGTACATTCCTGCGTCCGATCGACAGTAAAGACTGGCCCC 1055
QY 1321 CAABAAGCAGACCTGTAATGAAGATTTCCATGATCACCATCGTACCTATGGAAGTACTTA 1380
Db 1054 CAABAAGCAGACCTGTAATGAAGATTTCCATGATCACCATCGTACCTATGGAAGTACTTA 995
QY 1381 AGTAAATGATGTTCTTAAATGATGTTAATATAGCTTCATGAACCCCATTAACCTGGA 1440
Db 994 AGTAAATGATGTTCTTAAATGATGTTAATATAGCTTCATGAACCCCATTAACCTGGA 935
QY 1441 CACTTCTCTCAATCATGATGAGAGCTTTCGAGGCGCTCTTAGCCCTTTCCAT 1500
Db 934 CACTTCTCTCAATCATGATGAGAGCTTTCGAGGCGCTCTTAGCCCTTTCCAT 875
QY 1501 TTCCCTCTCGGCGGTGAGTCAGCGCTTAAAGCTCTCGAGCATCAGTGGGTTCAGTGT 1560
Db 874 TTCCCTCTCGGCGGTGAGTCAGCGCTTAAAGCTCTCGAGCATCAGTGGGTTCAGTGT 815
QY 1561 GGTATCGCGCACCCCAACCGCGCAAGCAACCGAGCGGTAAGTCTTGTGCAAGCAGC 1620
Db 814 GGTATCGCGCACCCCAACCGCGCAAGCAACCGAGCGGTAAGTCTTGTGCAAGCAGC 755
QY 1621 CGGAAAATGGTGTGCTGCAAGCAAACTTAAGCGCAACTATCTGAGCGCTTGTCTGA 1680
Db 754 CGGAAAATGGTGTGCTGCAAGCAAACTTAAGCGCAACTATCTGAGCGCTTGTCTGA 695
QY 1681 AATCCCTTAAACCATCGCATCAACAGATTCGCTATCCACATGTTTCTCCCGTGT 1740
Db 694 AATCCCTTAAACCATCGCATCAACAGATTCGCTATCCACATGTTTCTCCCGTGT 635
QY 1741 CAACAGATGAGTCTTTCGGGTGGAGCAACGCTACGCTGCGTTCGGAAGATGAAGCGCA 1800
Db 634 CAACAGATGAGTCTTTCGGGTGGAGCAACGCTACGCTGCGTTCGGAAGATGAAGCGCA 575
QY 1801 CACATATCTTCTGCTCGGCGTGGAGATGTTTTCAGGCGGTAAACCGTGAAGTAAATCC 1860
Db 574 CACATATCTTCTGCTCGGCGTGGAGATGTTTTCAGGCGGTAAACCGTGAAGTAAATCC 515
QY 1861 CGTGGCGGATGTAAGTAGTAGAACTTGGAAACCATCGGCACTTGGCCATTCGAAACCCC 1920
Db 514 CGTGGCGGATGTAAGTAGTAGAACTTGGAAACCATCGGCACTTGGCCATTCGAAACCCC 455
QY 1921 CTCATTCGCGGATGCTTACATGTTGATGGGAACTAGATTGGGCTGCGATGCGCGTCTT 1980
Db 454 CTCATTCGCGGATGCTTACATGTTGATGGGAACTAGATTGGGCTGCGATGCGCGTCTT 395
QY 1981 ACGCTTCGCTCCCAAGATGCTTCAAGACGTCGAGCTGAGCGGCGCTGATGCTGTC 2040
Db 394 ACGCTTCGCTCCCAAGATGCTTCAAGACGTCGAGCTGAGCGGCGCTGATGCTGTC 335
QY 2041 TGTGGGCGCGCAGCGCGTATCCATTGTCGCGGAAAGGTTTGTGTGAGGCAATTCG 2100
Db 334 TGTGGGCGCGCAGCGCGTATCCATTGTCGCGGAAAGGTTTGTGTGAGGCAATTCG 275
QY 2101 CCGAGGCTTGTGTGGGCACTTCTTCCGAAACCCAGCTGCTCCATGCTTAAAGCAGG 2160
Db 274 CCGAGGCTTGTGTGGGCACTTCTTCCGAAACCCAGCTGCTCCATGCTTAAAGCAGG 215
QY 2161 AGAGTATCTCTCTCGATGAGATACCCATTGACACCGATGATTTGGCAACGATGCG 2220
Db 214 AGAGTATCTCTCTCGATGAGATACCCATTGACACCGATGATTTGGCAACGATGCG 155
QY 2221 CCGTGAATCTAGATCTTCTAGCTAGCTACAGACCGCTGCTTGTGATGACGCAATCGAGG 2280
Db 154 CCGTGAATCTAGATCTTCTAGCTAGCTACAGACCGCTGCTTGTGATGACGCAATCGAGG 95
QY 2281 ATTGGGCGCTTAGTACTTCTGAAAAGGTTTCAGGGTTTTCATCTTCCGCCCGCAGGAA 2340

Db	1329095	TTCCCCCTCGCGCGTGAGTCTAGCGCGGTTAAAGCTCTCGAGCATCTACGTGGGTCTGAGTGT	1329154
QY	1561	GGTATCGCGCACCAACCGGCGCAAAAGCAACGAAGCGGGTGAAGTCTTTGTGCAAGCAGC	1620
Db	1329155	GGTATCGCGCACCAACCGGCGCAAAAGCAACGAAGCGGGTGAAGTCTTTGTGCAAGCAGC	1329214
QY	1621	GCGGAATAGTGTCTGCTCGACGAGAACTAAAGCGCAACTATCTGAGCGCCTCTCTGA	1680
Db	1329215	GCGGAATAGTGTCTGCTCGACGAGAACTAAAGCGCAACTATCTGAGCGCCTCTCTGA	1329274
QY	1681	AATCCCGTTAAACCATCGCCATCAACGAGAACTCGTATCCATCAATGTTTCTCCCGTGT	1740
Db	1329275	AATCCCGTTAAACCATCGCCATCAACGAGAACTCGTATCCATCAATGTTTCTCCCGTGT	1329334
QY	1741	CACGAGGTAGCTCTTGGGTGGAGCAACGCTCACGCTGGCTGTGGAGATGAAGCGCA	1800
Db	1329335	CACGAGGTAGCTCTTGGGTGGAGCAACGCTCACGCTGGCTGTGGAGATGAAGCGCA	1329394
QY	1801	CACATATCTCTGTCTCGCGCTGGAGATGTTTATAGAGCGGTAAACCGGTGAAGCTAATCC	1860
Db	1329395	CACATATCTCTGTCTCGCGCTGGAGATGTTTATAGAGCGGTAAACCGGTGAAGCTAATCC	1329454
QY	1861	CGTGGCGGATGTAAGTATGTAACCTTGGACCAATGGCGCACTGGGCCATTTGGAAACCC	1920
Db	1329455	CGTGGCGGATGTAAGTATGTAACCTTGGACCAATGGCGCACTGGGCCATTTGGAAACCC	1329514
QY	1921	CTCATTTGGCGATGCCCTACATGTTTGATGGGAACTAGATGGGCTGGCGATGCCGCTCT	1980
Db	1329515	CTCATTTGGCGATGCCCTACATGTTTGATGGGAACTAGATGGGCTGGCGATGCCGCTCT	1329574
QY	1981	ACGCTTCGGTCCAAAGATGTGCTTCAAGACCGTGAACCTGGACGGCGCGTGCATGGTCC	2040
Db	1329575	ACGCTTCGGTCCAAAGATGTGCTTCAAGACCGTGAACCTGGACGGCGCGTGCATGGTCC	1329634
QY	2041	TGTGGGGCGCAGGCGGTATCCATTTGCCGTGGCGGAAGTTTTGGTAGGCGCAATTGC	2100
Db	1329635	TGTGGGGCGCAGGCGGTATCCATTTGCCGTGGCGGAAGTTTTGGTAGGCGCAATTGC	1329694
QY	2101	CCGAGCGCTTGGTTGGGAGCTTCTTCCCGAATCCCAAGCTGCTCCCATGCTAAAGACGG	2160
Db	1329695	CCGAGCGCTTGGTTGGGAGCTTCTTCCCGAATCCCAAGCTGCTCCCATGCTAAAGACGG	1329754
QY	2161	AGAAGTGATCTCTCGATGAGATACCACTGACACACCGATGATTTGGCAACGATGGCG	2220
Db	1329755	AGAAGTGATCTCTCTCGATGAGATACCACTGACACACCGATGATTTGGCAACGATGGCG	1329814
QY	2221	CTTGGAATCTAGATCTCTAGCTPAGACTCAAGACCGCGTGTGTGATGAGCAATTCGAGG	2280
Db	1329815	CTTGGAATCTAGATCTCTAGCTPAGACTCAAGACCGCGTGTGTGATGAGCAATTCGAGG	1329874
QY	2281	ATTGGGGCTTAGTTACTCTTGAAAGGTTTCAGGGTTTTTCATCTTTCGCCCGCGAGGAA	2340
Db	1329875	ATTGGGGCTTAGTTACTCTTGAAAGGTTTCAGGGTTTTTCATCTTTCGCCCGCGAGGAA	1329934
QY	2341	TGGGCGAGGAGATGAACACTTCAGCAAAATGG	2374
Db	1329935	TTGGGCGAGGAGATGAACACTTCAGCAAAATGG	1329968

RESULT 3

```

RES001.3
US-09-746-660A-47
; Sequence 47, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
;

```

```

RESULT 4
US-09-738-626-3456
; Sequence 3456, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKTO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3456
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3456

Query Match      36.6%; Score 870; DB 9; Length 870;
Best Local Similarity 100.0%; Pred. No. 6e-280;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1421  ATGACCCCATTCACATCGACACTTCTCTCAATCATTCATGACGAGGAGCTTCGAGGC 1480
Db       1  ATGACCCCATTCACATCGACACTTCTCTCAATCATTCATGACGAGGAGCTTCGAGGC 60
QY      1481  GCCTCTCTAGCCCTTCTTCATTTCCCTCCCTCGGGGGTGAGTCACGGGCTTAAAGCTCTCGAG 1540

```


QY 1069 CTTGTTTGTCAATGGCGCTTTTCGCTGCCATGACGGCAAAACCAATACAGGTAAGCGATGCC 1128
 Db 291 CTTGTTTGTCAATGGCGCTTTTCGCTGCCATGACGGCAAAACCAATACAGGTAAGCGATGCC 232
 QY 1129 ACCCCAGCGCATTAATATCAGACGATCGCGCGGCAATGACAAAGATCAAGCCCAA 1188
 Db 231 ACCCCAGCGCATTAATATCAGACGATCGCGCGGCAATGACAAAGATCAAGCCCAA 172
 QY 1189 GTGTCGCGCGCATCAACAAAGAGCGTCAGAAATTAACACACGAGGAACCGCAATGAG 1248
 Db 171 GTGTCGCGCGCATCAACAAAGAGCGTCAGAAATTAACACACGAGGAACCGCAATGAG 112
 QY 1249 TCCCTTCGCGCTTAATTCCTTTGTTTAATCACCAGTCAATCTCGCGTCCGATGGAAGTAA 1308
 Db 111 TCCCTTCGCGCTTAATTCCTTTGTTTAATCACCAGTCAATCTCGCGTCCGATGGAAGTAA 52
 QY 1309 AGACTGCG 1359
 Db 51 AAGACTGCG 1

RESULT 7
 US-09-738-626-3455/c
 ; Sequence 3455, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: Patent in ver. 3.0
 ; SEQ ID NO 3455
 ; LENGTH: 708
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3455

Query Match 29.8%; Score 708; DB 9; Length 708;
 Best Local Similarity 100.0%; Pred. No. 1e-225;
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 652 ACCCATCAATCAGTTTGTATGCGCAATGCGGTCTCATCACTGCCACGACGATTGAT 711
 Db 708 ACCCATCAATCAGTTTGTATGCGCAATGCGGTCTCATCACTGCCACGACGATTGAT 649
 QY 712 CCAGCGCCACACCTTGGGCGTGGACGCGGCGTGCACATCTGCTGCGCGCAAAACCCAC 771
 Db 648 CCAGCGCCACACCTTGGGCGTGGACGCGGCGTGCACATCTGCTGCGCGCAAAACCCAC 589
 QY 772 CCAGCGGGAACACAGATCAGCTTGGCGGGAACCGCGCGCGGGAATATCCACGTCGCGT 831
 Db 588 CAGCGGGAACACAGATCAGCTTGGCGGGAACCGCGCGCGGGAATATCCACGTCGCGT 529
 QY 832 GTCCGCGGTATGCGCGCGACCGCGCGGATTAACACAAACGCTCCAAATAGCATTCGG 891
 Db 528 GTCCGCGGTATGCGCGCGACCGCGCGGATTAACACAAACGCTCCAAATAGCATTCGG 469

QY 892 GTTCAACAGGTGACGACGATGCCATCAATGGCGCTTTACCCAAACCGCGCTGCTTATC 951
 Db 468 GTTCAACAGGTGACGACGATGCCATCAATGGCGCTTTACCCAAACCGCGCTGCTTATC 409
 QY 952 GAGCGTCACTCCACCCGACCGCGGTGCGCGTGTCAAGTGGCCACCGCGCAACCGCCAA 1011
 Db 408 GAGCGTCACTCCACCCGACCGCGGTGCGCGTGTCAAGTGGCCACCGCGCAACCGCCAA 349
 QY 1012 AGCGGTGTCATCGCGGACCGGTTGTTCTTCTTCAATGATCTGTGGCGCTTCCACCTT 1071
 Db 348 AGCGGTGTCATCGCGGACCGGTTGTTCTTCTTCAATGATCTGTGGCGCTTCCACCTT 289
 QY 1072 GTTGTGTCATCGCGGCTTTTCGCTGCCATGAGGCAAAACCAATACAGGTAAGCGATGCC 1131
 Db 288 GTTGTGTCATCGCGGCTTTTCGCTGCCATGAGGCAAAACCAATACAGGTAAGCGATGCC 229
 QY 1132 CCAGCGCATATATCGAGACGATCGCGCGGCAATGGAACAAAGATCAACGCGCAAGT 1191
 Db 228 CCAGCGCATATATCGAGACGATCGCGCGGCAATGGAACAAAGATCAACGCGCAAGT 169
 QY 1192 GCGCGGATGAACAAAGACGCTCAGAAATTAACACACGAGGAACCGCAATGAGTCC 1251
 Db 168 GCGCGGATGAACAAAGACGCTCAGAAATTAACACACGAGGAACCGCAATGAGTCC 109
 QY 1252 TTGCGGCTTAATTCCTTTGTTTAAATCACCAGTACATTTCTGCGGTCCGATGGACAGTAAAG 1311
 Db 108 TTGCGGCTTAATTCCTTTGTTTAAATCACCAGTACATTTCTGCGGTCCGATGGACAGTAAAG 49
 QY 1312 ACTGCG 1359
 Db 48 ACTGCG 1

RESULT 8
 US-10-166-142-9/c
 ; Sequence 9, Application US/10166142
 ; Publication No. US20030124687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GUNJI, YOSHIYA
 ; APPLICANT: YASUEDA, HISASHI
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE OR L-ARGININE BY USING METHANOL
 ; TITLE OF INVENTION: ASSIMILATING BACTERIUM
 ; FILE REFERENCE: 223789US
 ; CURRENT APPLICATION NUMBER: US/10/166,142
 ; CURRENT FILING DATE: 2002-06-11
 ; PRIOR APPLICATION NUMBER: JP 2001-1777075
 ; PRIOR FILING DATE: 2001-06-12
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 9
 ; LENGTH: 712
 ; TYPE: DNA
 ; ORGANISM: Brevibacterium lactofermentum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(375)
 ; OTHER INFORMATION:
 US-10-166-142-9

Query Match 29.4%; Score 696.8; DB 14; Length 712;
 Best Local Similarity 99.6%; Pred. No. 5.7e-222;
 Matches 709; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 649 CTAACCCATCAATCAGTTTGTATGCGCAATGCGGTCTCATCACTGCCACGACGCTT 708
 Db 712 CTAACCCATCAATCAGTTTGTATGCGCAATGCGGTCTCATCACTGCCACGACGCTT 653
 QY 709 GATCCAGCGCGCACACCTTGGGCGTGGACGCGGCGTGCACATCTGCTGCGCGCAAAACC 768
 Db 652 GATCCAGCGCGCACACCTTGGGCGTGGACGCGGCGTGCACATCTGCTGCGCGCAAAACC 593
 QY 769 CACACGCGGGAACACAGATCAGGCTTTGCGCGCAACGCGCGCGGGAATATCCACGCTCC 828

Db 592 CACCAGCGGAGCAGATCAGGCTTCCGCGAGCGCGCGGAGAAATCCACGCTCC 533
 QY 829 GGTGTGCGGATTTGCGCGCGAGCGCGCGGATTAACAAACCGGTCCTCAATTC 888
 Db 532 GGTGTGCGGATTTGCGCGCGAGCGCGCGGATTAACAAACCGGTCCTCAATTC 473
 QY 889 CGGGTCAACAGAGTCAAGCAGATTCAGCAGATTCAGCAGATTCAGCAGATTC 948
 Db 472 CGGGTCAACAGAGTCAAGCAGATTCAGCAGATTCAGCAGATTCAGCAGATTC 413
 QY 949 ATCCAGCGCTCACTCCACCGCGACCGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1007
 Db 412 ATCCAGCGCTCACTCCACCGCGACCGCGGTGCGGTGCGGTGCGGTGCGGTGCG 353
 QY 1008 CAAAGCGGTGTCATCGCGCGAGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1067
 Db 352 CAAAGCGGTGTCATCGCGCGAGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 293
 QY 1068 CTTTGTGTCATCGCGCTCTTTCGTCGATGCGGCAACCAATCAAGGTAAAGCGATGC 1127
 Db 292 CTTTGTGTCATCGCGCTCTTTCGTCGATGCGGCAACCAATCAAGGTAAAGCGATGC 233
 QY 1128 CACCCAGCGCATTAATCGAGCAGATCGCGCGCGATCGCGCGCGATCGCGCGCG 1187
 Db 232 CACCCAGCGCATTAATCGAGCAGATCGCGCGCGATCGCGCGCGATCGCGCGCG 173
 QY 1188 AGTGCGCGCGATGAGCAAAAGAGCGTGAATTAACACAGAGAGAGACCGGATGA 1247
 Db 172 AGTGCGCGCGATGAGCAAAAGAGCGTGAATTAACACAGAGAGAGACCGGATGA 113
 QY 1248 GTCTTCGCGCTTAATTCCTTTTAAATCAGCAGTACATTCGCGGTCCGATGACAGTA 1307
 Db 112 GTCTTCGCGCTTAATTCCTTTTAAATCAGCAGTACATTCGCGGTCCGATGACAGTA 53
 QY 1308 AAAGACTGCG 1359
 Db 52 AAAGACTGCG 1

RESULT 9
 US-09-738-626-3454
 ; Sequence 3454, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738, 626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: Patent in ver. 3.0
 ; SEQ ID NO 3454
 ; LENGTH: 627
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3454

Query Match 26.4%; Score 627; DB 9; Length 627;
 Best Local Similarity 100.0%; Pred. No. 1.3e-198; Mismatches 0; Indels 0; Gaps 0;
 Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 ATGTACGCATTCGCTGACATTTGCTGCTGAAAAGGCTCTTTAGTGGGTATTCTTCC 82
 Db 1 ATGTACGCATTCGCTGACATTTGCTGCTGAAAAGGCTCTTTAGTGGGTATTCTTCC 60
 QY 83 TAGCGTCCAGAGCTCAGCGGAGGCGGCTGAGTTCATGCGGAGGAGGCTGCCGCTT 142
 Db 61 TAGCGTCCAGAGCTCAGCGGAGGCGGCTGAGTTCATGCGGAGGAGGCTGCCGCTT 120
 QY 143 CTGATTCATCAGCAAGCTATTCATCAATTAATGTTGGGTGAGGAAACCGGCGATGAC 202
 Db 121 CTGATTCATCAGCAAGCTATTCATCAATTAATGTTGGGTGAGGAAACCGGCGATGAC 180
 QY 203 GGTGAGAACTTTGTGAGTCAAGTCCAAATGCTTCTTGGGTCATTCGCTTTCTACCA 262
 Db 181 GGTGAGAACTTTGTGAGTCAAGTCCAAATGCTTCTTGGGTCATTCGCTTTCTACCA 240
 QY 263 CTTTGGCAGGCGCTCTCAGGCAAAATATCTCGATGGAATTCAGAGGTTCCCGGCGC 322
 Db 241 CTTTGGCAGGCGCTCTCAGGCAAAATATCTCGATGGAATTCAGAGGTTCCCGGCGC 300
 QY 323 AGCCAGGCTAAGTCCCTGCTGAGGCGCATTTGAACGCTGAACAAATATGATATGCTCCGC 382
 Db 301 AGCCAGGCTAAGTCCCTGCTGAGGCGCATTTGAACGCTGAACAAATATGATATGCTCCGC 360
 QY 383 AAGCTCAATGACATCGCCAGGAAACCGGCGCGATCTTCCGCGAGATGCGCTTGCATGG 442
 Db 361 AAGCTCAATGACATCGCCAGGAAACCGGCGCGATCTTCCGCGAGATGCGCTTGCATGG 420
 QY 443 GTGCTGCGGAGAGAGAGTACGCGCGGATACGCGGATACGCGGATGCGCTTGCATGGCT 502
 Db 421 GTGCTGCGGAGAGAGAGTACGCGCGGATACGCGGATGCGCTTGCATGGCT 480
 QY 503 TCGTCAAGTTCAGAGCTGGAACAGCGCTTGAATCACTCAACAACTTGGAGTTTCTGAC 562
 Db 481 TCGTCAAGTTCAGAGCTGGAACAGCGCTTGAATCACTCAACAACTTGGAGTTTCTGAC 540
 QY 563 GCGAGTTGAGGCGATGATGAGATTTCCACAGCGCGGATCAACATTTGGGCGAG 622
 Db 541 GCGAGTTGAGGCGATGATGAGATTTCCACAGCGCGGATCAACATTTGGGCGAG 600
 QY 623 GCCACCGATTTCCAAAACCGCGGAAAC 649
 Db 601 GCCACCGATTTCCAAAACCGCGGAAAC 627
 RESULT 10
 US-10-627-476-669
 ; Sequence 669, Application US/10627476
 ; Publication No. US20040030116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompejus, Mark
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schoder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Haberhauer, Gregor
 ; TITLE OF INVENTION: CORINEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 ; INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
 ; FILE REFERENCE: BGI-125CPCN
 ; CURRENT APPLICATION NUMBER: US/10/627,476
 ; CURRENT FILING DATE: 2003-07-25
 ; PRIOR APPLICATION NUMBER: 09/602,787
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: USSN 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: DE 19931454.3
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931478.0
 ; PRIOR FILING DATE: 1999-07-08

;; PRIOR APPLICATION NUMBER: DE 19931563.9
;; PRIOR FILING DATE: 1999-07-08
;; PRIOR APPLICATION NUMBER: DE 19932122.1
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: DE 19932124.8
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: DE 19932125.6
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: DE 19932128.0
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: DE 19932180.9
;; PRIOR FILING DATE: 1999-07-09
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 678
;; SEQ ID NO 669
;; LENGTH: 993
;; TYPE: DNA
;; ORGANISM: Corynebacterium glutamicum
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (101)..(970)
;; OTHER INFORMATION: RKN03164
US-10-627-476-669

Query Match 20.4%; Score 485.2; DB 12; Length 993;
Best Local Similarity 99.0%; Pred. No. 5.8e-151;
Matches 499; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1 AGATACCTCTTGGAGAACCATGTACGATTCGGTGACATTTGCGTCTGGAAGGC 60
Db 490 AGATACCTCTTGGAGAACCATGTACGATTCGGTGACATTTGCGTCTGGAAGGC 549

QY 61 TCTTTACGTGGTATTTCTTCTACGCTCAGAGCTCAGCGAGCGCGCTGAGTTCA 120
Db 550 TCTTTACGTGGTATTTCTTCTACGCTCAGAGCTCAGCGAGCGCGCTGAGTTCA 609

QY 121 GCGGAGAGAGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTG 180
Db 610 GCGGAGAGAGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTG 669

QY 181 GGTGAGAACCGGGGATGAGGTGAGACTTTGTCAGTCACTGCCAACAATGCTCT 240
Db 670 GGTGAGAACCGGGGATGAGGTGAGACTTTGTCAGTCACTGCCAACAATGCTCT 729

QY 241 TGGCGTCAATGCTTTCTACCACTTGGCGAGGCGCTGCTCAGCAACAATTCGATGG 300
Db 730 TGGCGTCAATGCTTTCTACCACTTGGCGAGGCGCTGCTCAGCAACAATTCGATGG 789

QY 301 AATTCAGAGGGTTCGCGCGCAGCCAGGCTAAGTCCCTGCTCAGGCAATGTTGAAGT 360
Db 790 AATTCAGAGGGTTCGCGCGCAGCCAGGCTAAGTCCCTGCTCAGGCAATGTTGAAGT 849

QY 361 GAACAAATTCATGATGCTCCGCAAGCTCAATGATCGCCAGCAAGCGGCGAGTCACT 420
Db 850 GAACAAATTCATGATGCTCCGCAAGCTCAATGATCGCCAGCAAGCGGCGAGTCACT 909

QY 421 TCGCAGATGCGGCTTGCATGAGGCTGCTCGCGAGCAAGGAGTACGCGGC--GGATACC 478
Db 910 TCGCAGATGCGGCTTGCATGAGGCTGCTCGCGAGCAAGGAGTACGCGCGGATACC 969

QY 479 GTGACAGTGCATGATGCTGCT 502
Db 970 GTGACAGTGCATGATGCTGCT 993

RESULT 11
US-10-282-122A-17521
; Sequence 17521, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Kari
;; APPLICANT: Zyskind, Judith
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John
;; APPLICANT: Carr, Grant
;; APPLICANT: Yamamoto, Robert
;; APPLICANT: Forsyth, R.
;; APPLICANT: Xu, H.
;; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;; FILE REFERENCE: ELITRA.034A
;; CURRENT APPLICATION NUMBER: US/10/282,122A
;; CURRENT FILING DATE: 2003-02-20
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SEQ ID NO 17521
;; SOFTWARE: PatentIn version 3.1
;; LENGTH: 1095
;; TYPE: DNA
;; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-17521

Query Match 12.9%; Score 305.8; DB 12; Length 1095;
Best Local Similarity 67.1%; Pred. No. 8e-91;
Matches 433; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

QY 1 AGATACCTCTTGGAGAACCATGTACGATTCGGTGACATTTGCGTCTGGAAGGC 60
Db 450 AGATACCTCTTGGAGAACCATGTACGATTCGGTGACATTTGCGTCTGGAAGGC 509

QY 61 TCTTTACGTGGTATTTCTTCTACGCTCAGAGCTCAGCGAGGCGGCTGAGTTCA 120
Db 510 TCGCTATGTTGGATTTCTTCTTATGGACCTGAATAACGAGAGATCTCGAATTC 569

QY 121 GCGGAGAGAGGCTGCGCGCTTCTGATTCATCAGCAAGCTATTCATCAATTAATCGTTG 180
Db 570 AGCTGTGAGGCTGCGCGCTTCTGATTCATCAACGAGCTATTCATTTGAACGCTG 629

QY 181 GGTGAGGAGAACCGGGCGATGACGCTGAGAACTTTGTCAGTCACTGCCAACAATGCTCT 240
Db 630 GGTGAGGAGAACCGGGCGATGAGTGGCGAGAACCTTTAGAGTCTCGCGGATTAACGACT 689

QY 241 TGGCGTCAATGCTTTCTACCACTTGGCGAGGCGCTCAGCAACAATATCTCGATGG 300
Db 690 GGTGTTATTCATTTTCGCCACTTGCACAGGCGCTTTTAATGACAGTATCTCGAGGG 749

QY 301 AATTCAGAGGGTTCCCGCGCAGCCAGCGGTAAGTCCCTGCTGAGGCGCATGTTGAAGCT 360
Db 750 CGTGCCAGTGGATTCAACGCGCAGCAGCAGGGAAGTCTTTGGGCAAGAGATGCTCAAGCG 809

QY 361 GAACAAATTCATGATGCTCGCAAGCTCAATGATCGCCAGCAAGCGGCGAGTCACT 420
Db 810 TAAAAATCTTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869

Qy	421	TGGCAGATGGCGCTTGCATATGGGTCTGCGCGAGCAAGAGAGTACGGCGCGGATACGT	480
Db	870	GGCACAGCTGGCTATTGGCTGGGTGCTGCGGGAAACAAGCGATTATTGGGGCCACACCGT	929
Qy	481	GACCACTGCATTTGATTTGGTGTCTCGTTCAGCTTCAGCACTCGACAACAGCTTCGATCTACT	540
Db	930	AACCACTGCGTTGATCGAGAGCATCGTCGGTAGCGAGTTGGATCAAAATCTCGGAGCTCT	989
Qy	541	CAACAACCTTGAGTCTTCTGACGCGGAGTTGAGCGCATCGATGAGATTTCCCAACGACGC	600
Db	990	TAAATATCTTGAGTTCAGCGTTGAGGAACGCAATTTATTTGACATGTTGCCAAGGACGC	1049
Qy	601	CGGCATCAACAATTTGGGCGAAGGCCACCGGATTTCAAAACCCGCA	645
Db	1050	TGCAATTTAATTTGGCGGGTGTATCTGCTTCGAAGTTTACGA	1094

RESULT 12

US-10-282-122A-17523
Sequence 17523, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Hazelbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Yu, H.

Query Match 12.8%; Score 302.8; DB 12; Length 879;
Best Local Similarity 59.2%; Pred. No. 7e-90;
Matches 517; Conservative 0; Mismatches 357; Indels 0; Gaps 0;
ov 1421 ATGARACCCCATTCACCTGCACATTTGCTCTCAATCATTTGATGAGGCGAGCTTCGAGGC 1480

Db	1	ATGAATCCGGTCTCATTTTAAAGACCTTGTCTGCGCAATATGACGAAGGCGCTTCGAAGAT	60
Qy	1481	GCCTCCTTTAGCCCTTTCCATTTTCCCTCTCGGCGGTGAGTCAGCGGTTAAAGCTCTCGAG	1540
Db	61	CGGGCTTTTGCATTAGGAATTTCTCTTCAGCAGTAAGTCAGAGGATTAAGGCGTTGGAG	120
Qy	1541	CATCAGTGGGTGAGTGTGTGATTCGGCCACCCACCGGCCAAAGCAACCGAAGCGGGT	1600
Db	121	CGGGAGACTGTCGATAGTGATGTCACGCGGTACGTACACAGTCACCTGCACTGCACTGATGCGCGC	180
Qy	1601	GAAGTCTCTTGTCAAGCAGCGCCGAAAAATGTGTGTCTGCAAGCAGAAAATAAGACGCGAA	1660
Db	181	GAGGTACTGTGCAATCGGCACGACGTATGGCACTTCTGCGGCCGGAACAAACGCTCAG	240
Qy	1661	CTATCTGACGCGCTTGTGAAATCCCGTTAAGCATCGCCATCGACGACAGATTCGCTATCC	1720
Db	241	TTGCGGGGAAGAAATTGAGCGGATTCGATTTGTCTGTAGCTGTCAACGCGAGATTTCTTGCC	300
Qy	1721	ACATGTTTTCTCCCGTGTTCACGAGGTAGCTTCTTTGGGTGGAGCAACGCTCACGGTG	1780
Db	301	ACATGTTTTGCCCGGTGATGGCAAGTGTGGCGGAGATGATAAAGCCACATTCGATATG	360
Qy	1781	CGTTTGAAGATCAAGCGCAACATATTCCTTGTCTCGCGGTGGAGATGTTTTAGAGCG	1840
Db	361	CGTATCGAGGACGAGTCGCAATCGTTGTCTGTGTGTGCGTGGCGGAGATGTCCTTGGCGCT	420
Qy	1841	GTAAACCGGTGAAGCTAATCCCGTGGCGGCAATGTGAAGTAGTAGAACTTTGGAAACATAGGC	1900
Db	421	GTAAACGGGGAAAGACACCTGTATCTGTTTGGACTCGATAGCGTTAGTGCATTCGGA	480
Qy	1901	CATTTGGCCATTTGCAACACCCCTCATTTGCGGAGTGCCTATATGGTTGATGGGAACTAGAT	1960
Db	481	TATTTTGCAGTGGCAAGTCCGAAGCTGTTTGGATCGCTACACCATAGACGGTGTGTTGAT	540
Qy	1961	TGGGTGCGGATGCCCGCTTTACGCTTCGGTCCGATCCAAAGATGCTTCAAGACCGTGACCTG	2020
Db	541	TGGGACCGGATTCGAGTATTCGATTTGGCCCTTCGGATGGGCTTCAGACCGTGATATT	600
Qy	2021	GACGGGCGCGTCGATGGTCTCTGTGGGCGCAGGCGCGTATCCATTTGCCGTGGCGGAA	2080
Db	601	AAGCGTCTCTTGGTGAAGTTCCGGCAGGCGCGAAGTGTAAATGAGATCCCTTCATCGGAG	660
Qy	2081	GGTTTGTGTAGGCAATTTGCGGAGGCTTGGTTGGGGACCTTCTCCGAAACCCCAAGCT	2140
Db	661	GCATTCATGGAGCGGCCCGTGTGGGATTTGGGCTTGGGCAATGCTCCCCGATATTCAAGCG	720
Qy	2141	GCTCCCATGCTAAAGCAGGAGAAATGATCTCTCTTCGATGAGATACCATTTGACACACCG	2200
Db	721	TTACCGCTTTTGGATTTCTGTTGAGTAGTGTGCTAGATAGCGAGTGCACGAAGTTCCG	780
Qy	2201	ATGTAATGGCAACGATGGCGCTCGAATCTAGATCTCTAGCTAGACTCACAGACGCGCTC	2260
Db	781	CTGTATTTGGACGATTTGGCGGTTTGAATCTTCAGGCCCTTAGAGCGGTTGACTCAAGCGGTA	840
Qy	2261	GTTTGTGACGCAATCGAGGGATTCGCGCCTTAGT	2294
Db	841	GTGGATGCAACCTTCAGGGACTACATCACTGT	874

```

RESULT 13
US-10-369-493-35374
; Sequence 35374, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiaofeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCES: 38-10(52052)B

```


; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35374
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-35374

Query Match 6.9%; Score 164.4; DB 15; Length 1026;
Best Local Similarity 57.2%; Pred. No. 1.8e-43;
Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;
QY 3 ATACTCTTTGGAAGAAACATGACGATTCGTCGACATGTTGGTCTGGAAGGCTC 62
DB 428 ACACACCGCTTGAGAAACCTCGCGCGCTGGACAGATCGTGGCTCCGGCAAGGCGC 487
QY 63 TTTACGTTGGGATTTCTTCTTACGCTCCAGAGCTCACAGCGAGCGGCTGAGTTCAATGG 122
DB 488 TCTATGTCGGCATCTCTCTTACACTGAGCGGACCGCGAGCGCGCTATCTCTGA 547
QY 123 CGGAGAGGGCTGCGCGCTCTGATTCATCAGCCAGCTATTCATTAATCGTTGGG 182
DB 548 AGGATCTCGGCAAGCGCTGATCATCCACAGCGGAGCTATTCGATGATCAACCGCTGA 607
QY 183 TGGAGAACCGGCGATGAGCTGAGAACTTTGTCAGTCAGCTGCCAACAATGCTTGG 242
DB 608 TCGAGGAAGCGCTTTGTCG-----ATACGCTGGAAGAACTGGGTATCG 652
QY 243 GCGTCATTTCTTCTACCACTTGGCAGGCGCTGCTCACGAGCAAAATATCTCGATGAA 302
DB 653 GCTCCATGCTTTTTCGCGCTGGCGAGGCGATGCTGACAGCAAAATATCTGGCGGTG 712
QY 303 TTCCAGAGGTTCCGCGCGCAGCAAGGTAAGTCCCTGTCTGAGGGCAGTTGAAAGTGA 362
DB 713 TGCAGATGGCAGCGCTGCTCACAGAGCAAGTCACTCAACCGCGCTTCTCAACGAGC 772
QY 363 ACAATATGATGTCCTGCGAGCTCAATGACATGCCAGACGCGGCGAGTCACTTG 422
DB 773 GCAATGTCGAAACATCCGCGCTGAAACAGCATTCGCGAGCGGTGGCAGAGCTGG 832
QY 423 CGCAGATGGCGCTTGCATGGGTGCTGCGGAGCAAGAGAGTACGGCGGAGTACCGTGA 482
DB 833 CGCAGATGGCAATTCCTGGGTCTGCGCGCGCG-----CATTA 874
QY 483 CCAGTGCATTTGATGTTGCTTCTGATGAGCAGCTGGAACAAGCCTTGAATCACTCA 542
DB 875 CTTAGCATTTGATGGCGAAGCGGTGCGAACAAGGTCGAGGACTGCGTGAAAGCACTCG 934
QY 543 ACAACTTGGAGTTTCTGAGCGCGAGTTGGAGCGGATGATGAGATTTCCCAAGCGCG 602
DB 935 ATATGCGGAGTCTCTTACCGAGGAGCTGGCCGAAATCGACCGTTACGCCAAGGATGCG 994
QY 603 GCATCAACATTTGGCGAAGGCGCACCGA 630
DB 995 ATATCAACCTCTGGCGAAATCTGCCGA 1022

RESULT 14

US-10-369-493-38826
; Sequence 38826, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38826
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38826

Query Match 6.9%; Score 164.4; DB 15; Length 1026;
Best Local Similarity 57.2%; Pred. No. 1.8e-43;
Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;
QY 3 ATACTCTTTGGAAGAAACATGACGATTCGTCGACATGTTGGTCTGGAAGGCTC 62
DB 428 ACACACCGCTTGAGAAACCTCGCGCGCTGGACAGATCGTGGCTCCGGCAAGGCGC 487
QY 63 TTTACGTTGGGATTTCTTCTTACGCTCCAGAGCTCACAGCGAGCGGCTGAGTTCAATGG 122
DB 488 TCTATGTCGGCATCTCTCTTACACTGAGCGGACCGCGAGCGCGCTATCTCTGA 547
QY 123 CGGAGAGGGCTGCGCGCTTCTGATTCATCAGCCAGCTATTCATTAATCGTTGGG 182
DB 548 AGGATCTCGGCAAGCGCTGATCATCCACAGCGGAGCTATTCGATGATCAACCGCTGA 607
QY 183 TGGAGAACCGGCGATGAGCTGAGAACTTTGTCAGTCAGCTGCCAACAATGCTTGG 242
DB 608 TCGAGGAAGCGCTTTGTCG-----ATACGCTGGAAGAACTGGGTATCG 652
QY 243 GCGTCATTTCTTCTACCACTTGGCAGGCGCTGCTCACGAGCAAAATATCTCGATGAA 302
DB 653 GCTCCATGCTTTTTCGCGCTGGCGAGGCGATGCTGACAGCAAAATATCTGGCGGTG 712
QY 303 TTCCAGAGGTTCCGCGCGCAGCAAGGTAAGTCCCTGTCTGAGGGCAGTTGAAAGTGA 362
DB 713 TGCAGATGGCAGCGCTGCTCACAGAGCAAGTCACTCAACCGCGCTTCTCAACGAGC 772
QY 363 ACAATATGATGTCCTGCGAGCTCAATGACATGCCAGACGCGGCGAGTCACTTG 422
DB 773 GCAATGTCGAAACATCCGCGCTGAAACAGCATTCGCGAGCGGTGGCAGAGCTGG 832
QY 423 CGCAGATGGCGCTTGCATGGGTGCTGCGGAGCAAGAGAGTACGGCGGAGTACCGTGA 482
DB 833 CGCAGATGGCAATTCCTGGGTCTGCGCGCGCG-----CATTA 874
QY 483 CCAGTGCATTTGATGTTGCTTCTGATGAGCAGCTGGAACAAGCCTTGAATCACTCA 542
DB 875 CTTAGCATTTGATGGCGAAGCGGTGCGAACAAGGTCGAGGACTGCGTGAAAGCACTCG 934
QY 543 ACAACTTGGAGTTTCTGAGCGCGAGTTGGAGCGGATGATGAGATTTCCCAAGCGCG 602
DB 935 ATATGCGGAGTCTCTTACCGAGGAGCTGGCCGAAATCGACCGTTACGCCAAGGATGCG 994
QY 603 GCATCAACATTTGGCGAAGGCGCACCGA 630
DB 995 ATATCAACCTCTGGCGAAATCTGCCGA 1022

RESULT 15

US-10-369-493-38124
; Sequence 38124, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

```
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38124
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38124

Query Match          6.9%; Score 164.4; DB 15; Length 1028;
Best Local Similarity 57.2%; Pred. No. 1.e-43;
Matches 359; Conservative 0; Mismatches 236; Indels 33; Gaps 2;

QY 3 ATATCTCTTTGGAAGAACCATGTACGCAATTGCGTGACATTTGTCGCTGCGAAAGGCTC 62
Db 427 ACACACCGCTTCAGGAAACCTGCGGCGCTGGACACAGATCGTGGCTCCGGCAAGGCGC 486

QY 63 TTTAGTGGGTATTTCTCTACGTTCCAGAGCTTCAGCGAGGCGGCTGAGTTTCATGS 122
Db 487 TCTATGTGCGGCATCTCTCTCAACTCGAAGCGCACCCGCGAGGCGCGCTATCTCTGA 546

QY 123 CGAGGAGGCGTCCCGCTTCTGATTCATCAGCCAAAGTATTCATTAATCGTTGGG 182
Db 547 AGGATCTCGGACGCGCTGCAATCATCCACGCGAGCTATTCGATGATCAACGCTGGA 606

QY 183 TGGAGGAACCGGCGGATGACGGTGAGAACTTGTGTGAGTCAGCTGCCAAATGTTG 242
Db 607 TCGAGGAAGACGGTCTTGTGCG-----ATACGCTGGAAGAACTGGGTATCG 651

QY 243 CGGTCAATGCTTCTCACCATTGCGGAGGCGCTGCTCAGGACAAATATCTCGATGGA 302
Db 652 GCTCCATCGTCTTTGCGCGCTGGCGGAGGCGATGCTGACGACGAAATATCTGGGCGGTG 711

QY 303 TTCCAGAGGCTTCCGCGCCAGCCAGGCTAAGTCCCTGCTGTGAGGGCATGTTGAACGTGA 362
Db 712 TGCCGGATGCGACCGCTGCTCAGAGGCAAGTCACTCAACCGGCTTCTCTCAACGAGC 771

QY 363 ACAATATTGATATGTCGCAAGCTCAATGACATCGCCAGGAAACCGGCGAGTCACTTG 422
Db 772 GCAATGTGGAAGAACATCCGCGCGCTGAACAGCAATTTGCCGAGCGCGCTGGCCAGAGCGTGG 831

QY 423 CGCAGATGCGGCTTGCATGGGTGCTGCGGACAGGAGAGTACGGCGGCGATACCGTGA 482
Db 832 CGCAGATGCGCAATTGCTTGGGTCTTGGCGGCGGCGCG-----CATTA 873

QY 483 CCAGTGCATTGATTGCTGCTTCGTGAGTTCGAGCAGCTGGACAAACAGCCTTGATTCACTCA 542
Db 874 CCTCAGCATTTGATTGGCGCAAGCCGTGTCGACAGCTCGAGGACTGGGTGAAGCACTCG 933

QY 543 ACAACTTGGAGTTTCTGACGCGAGTTGGAGCGGATCGATGAGATTTCCACGACCGCG 602
Db 934 ATAATGCCGAGTTCTCTACCGAGGAGCTGGCCGAAATCGACCGGTTACGCCAAGGATCGG 993

QY 603 GCATCAACATTTTGGCGAGGCGCACCGA 630
Db 994 ATATCAACCTCTGGGCAAAATCTGCCGA 1021
```

Search completed: March 17, 2004, 09:47:27
Job time : 866.071 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 13:43:00 ; Search time 6528.5 Seconds
(without alignments)
10858.975 Million cell updates/sec

Title: US-09-105-117K-3
Perfect score: 2374

Sequence: 1 agatactccttggagaaga.....gtaacaccttcagcaaatgg 2374

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_gss_hum:*
18:	em_gss_inv:*
19:	em_gss_pin:*
20:	em_gss_vrt:*
21:	em_gss_fun:*
22:	em_gss_mam:*
23:	em_gss_mus:*
24:	em_gss_pro:*
25:	em_gss_rod:*
26:	em_gss_phg:*
27:	em_gss_vrl:*
28:	gb_gss1:*
29:	gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	84.6	3.6	657	13	BQ155035 NF075E081
C 2	73.2	3.1	3237	28	BH770958 LLMGtag68
3	72	3.0	860	28	AF075981 AF075981
4	60.4	2.5	830	29	AL151258 Anopheles

C	5	56.8	2.4	1034	28	BZ554294
	6	54.2	2.3	897	28	BZ560550
	7	51.4	2.2	1620	28	BZ568946
	8	48.8	2.1	1101	29	CNS0175T
	9	48.6	2.0	833	28	BZ556963
C	10	48.6	2.0	925	29	CNS0091P
C	11	47.6	2.0	852	13	BX393687
	12	46	1.9	614	14	CA830247
C	13	46	1.9	720	14	CD881613
	14	44.4	1.9	620	14	CA829022
	15	44.4	1.9	644	13	BQ704105
C	16	44.4	1.9	664	29	CNS03K4J
C	17	44.4	1.9	1201	13	BX381961
	18	43.6	1.8	721	28	BZ562255
	19	43.2	1.8	1395	28	BZ574094
C	20	42.8	1.8	532	10	AW927407
	21	42.8	1.8	538	14	CD861975
C	22	42.4	1.8	699	13	BX424825
C	23	42.2	1.8	885	13	BX425603
	24	42	1.8	457	13	BY246241
	25	42	1.8	1201	13	BX381961
	26	41.8	1.8	947	29	CNS077MV
C	27	41.8	1.8	1083	29	CNS079WN
	28	41.6	1.8	432	14	CB639421
	29	41.6	1.8	558	12	BM712036
C	30	41.6	1.8	588	13	BU220119
	31	41.6	1.8	772	14	CB668778
C	32	41.6	1.8	803	14	CB669972
C	33	41.6	1.8	813	14	CB684812
C	34	41.6	1.8	1201	13	BX356664
C	35	41.6	1.8	1201	13	BX361080
C	36	41.2	1.7	910	29	CNS0060N
	37	41	1.7	617	12	BM729055
	38	41	1.7	770	10	BE898495
	39	41	1.7	861	10	BE740880
	40	40.8	1.7	605	13	CA127255
	41	40.8	1.7	614	13	CA141725
	42	40.8	1.7	619	14	CA192339
	43	40.8	1.7	619	14	CA238619
	44	40.8	1.7	950	11	AV104298
C	45	40.6	1.7	436	28	AQ405681

ALIGNMENTS

RESULT 1
BQ155035/c 657 bp mRNA linear EST 24-APR-2002
LOCUS NF075E081R1F1067 Irradiated Medicago truncatula cDNA clone
DEFINITION NF075E081R 5', mRNA sequence.

ACCESSION BQ155035
VERSION BQ155035.1 GI:20292094
KEYWORDS EST:
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 657)
AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation

TITLE Medicago truncatula irradiated library
JOURNAL Unpublished (2001)
COMMENT Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org

Insert Length: 657 Std Error: 0.00
 Plate: 075 row: E column: 08
 Seq primer: TCACAGGAACAGCTATGAC.

FEATURES

Location/Qualifiers
 1..657
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"
 /clone="NF075E081R"
 /tissue_type="seedlings"
 /dev_stage="seedling"
 /clone_lib="Irradiated"
 /note="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 3.6%; Score 84.6; DB 13; Length 657;
 Best Local Similarity 48.2%; Pred. No. 3.4e-11;
 Matches 237; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 1420 CATGAACCCCATTCACCTGGACACCTTGTCTCTCAATGATGATGAGCGAGCTTCGAAGG 1479
 DB 605 CCNCGACTACAGACCATACAGGCACTGGATCGGTGATACGTGAACGAGATTGAGCG 546

QY 1480 CGCTCTTTAGCCCTTTCCATTTCCCTCGCGGTGAGTCAGCGGTTAAAGCTCTCGA 1539
 DB 545 CGCGGCAAAAGCTGTGCATTTACACAACTCAGCGGTCTCACAGGCGCATTAAGCACTGGA 486

QY 1540 GCATCACTGGGTGAGTGTGGTATCGGCACCCCAACCGCCAAAGCAGACGAGCGGG 1599
 DB 485 AAATANGTTCGGGCGAGCGCGTGTGGTGGTACCGTACCGCGCGCCGACGGAACAGG 426

QY 1600 TGAAGTCTCTGTGCAAGCGCGGAAATGTTGTGTGCAAGCAGAACTAAAGCGCA 1659
 DB 425 GCNAAAATCTGTGCACCTGCTCGCGAGTGGAGTGTCTGGNAGAGAGTGGCTGGCGA 366

QY 1660 ACTATCTGGAGCGCTTGCTGAATAATCCCGTTAAACATCGCCATCAACGAGATTCGCTATC 1719
 DB 365 TGAACAAACCGGTTGCACTCCGCTGCTCTTCACTGGCGGTCAACGCCGACAGTCTGGC 306

QY 1720 CACATGTTTTCTCCCGTGTTCACAGAGTAGCTTCTTGGGTGGAGCAAGCTACAGCT 1779
 DB 305 GAGGTGTTGTTCTTGGTACTGGCTCGTGTGTTGCTGATTCGCTATTCGCTCACTT 246

QY 1780 GCGCTTGAAGATGAAGCGCACACATTTCTTCTGCTCGCGCTGGAGATGTTTTAGGAGC 1839
 DB 245 GCAGGTAGAGATGAACACCGCACTCAGGAACGTCTGCCCGCGCGAGTGTGGCGGC 186

QY 1840 GGTAAACCGCTGAAGCTAATCCCGTGGCGGATGTGAAGTAGTAGAATCTTGAACACATGCG 1899
 DB 185 GGTGAGTATTCAACATCAGGCGCTCCGAGTGTGTCTTGTTCGATAAATCTTGGTGGCTGA 126

QY 1900 CCATTGGCCAT 1911
 DB 125 CTATCTGTTCTG 114

RESULT 2

BH770958/c 3237 bp DNA linear GSS 01-MAY-2002
 LOCUS BH770958
 DEFINITION LLMGtag686 MG1363 Random Sequence Tag Library Lactococcus lactis

subsp. cremoris genomic, genomic survey sequence.

BH770958

VERSION BH770958.1 GI:20373915

KEYWORDS GSS.

SOURCE Lactococcus lactis subsp. cremoris

ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Lactococcus.

REFERENCE 1 (bases 1 to 3237)

AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.

TITLE Studies of genomes of dairy bacteria Lactococcus lactis

JOURNAL Sci. Aliments (2002) In press

COMMENT Contact: Sorokin A

Genetique Microbienne

INRA

CEJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16

Fax: 33 1 34 65 25 21

Email: sorokine@jouy.inra.fr

best homologue in strain IL1403 is nadR (98%)

Class: shotgun

High quality sequence start: 30

High quality sequence stop: 3207.

FEATURES

source

1..3237

/organism="Lactococcus lactis subsp. cremoris"

/mol_type="genomic DNA"

/strain="MG1363"

/sub_species="cremoris"

/db_xref="taxon:1359"

/clone_lib="MG1363 Random Sequence Tag Library"

/note="Vector: PSGM02; Site 1: SmaI; Library of

chromosomal fragments of L.lactis strain MG1363 was

prepared by partial AluI digestion or by sonication."

ORIGIN

Query Match 3.1%; Score 73.2; DB 28; Length 3237;
 Best Local Similarity 49.8%; Pred. No. 9.5e-08;
 Matches 287; Conservative 0; Mismatches 263; Indels 26; Gaps 3;

QY 1 AGATCTCTCTTTGGAAGAACCATGTACGATTCGCTGACATTTGTCGCTCGAAGGC 60
 DB 2511 AATACCCCTATTGGAAGAACCATGCGCGCTTTGAAACTGCTCGATAGTGGAAAGC 2452

QY 61 TCTTTACGTGGTATTTCTTCTTACGCTCCAGAGCTCACAGCGAGCGGCTGAGTTCA 120
 DB 2451 ACTTTATGTGGTGTGTCTTAATTTCTTCTGCAAGAAACCGAAGCTGCGGTCTTAGCGC 2392

QY 121 GCGGAGAGGGCTGCGCGCTCTGATTCATCAGCCAGCTATTCATCATTAATCGTTG 180
 DB 2391 TGAAGATTTAGGCTTTTAAACTCTTGATTCATCAGCCCTCGTTATCAATGTTAGATCGATG 2332

QY 181 GGTGAGGAACCGGCGCATGACGCTGAGAACTTGTTCAGTCAGCTGCCAACAATGGTCT 240
 DB 2331 GATTGAGATGATTTACAGAACTTTGACAGAGGGG-----GGAT 2289

QY 241 TGGCGTCAATTTCTTCTACCACTTGCAGCGGCTGCTTACGAGCAAAATATCTCGATGG 300
 DB 2288 AGGAACAATTTGCTTTAAGCTTTGATCAAGGACTCTTAAACAGGAAATATTTGTCATGG 2229

QY 301 AATTCAGAGGGTTC-----CGCGCCAGCGAGGTAGTCCCTGTCTGAGGGCATGTT 354
 DB 2228 AATTCCTGAAAATTCGCGAATCGGTGATCTCTCATTTATGCGACTTTTGCATGATGACAGTCT 2169

QY 355 GAAAGTGAACATATTCATATGTCGCAAGCTTCATGATCGCCAGGAACGCGGCA 414
 DB 2168 TACACAGAAAGATTGGAACAAGTTCAAGCCCTTATGATTTAGTCTAAAGTCTGTGACA 2109

QY 415 GTCACTTGGCAGATGCGCTTGCATGGGTGCTGCGGAGCAAGGAGAGTACGCGCGGA 474
 DB 2108 ATCTCTAGCTCAANTGCAATAAGCTTGGGTTTTTACG---GGAAGAGCTGCAAAAGTTCA 2052

QY 475 TACGCTGACCAGTCGATTTGTTGGTCTTCTGCTCAGTTGAGCAGCTGGCAACAGCCCTTGA 534

```

Db      2051 AGCTGTTACATCAGCTTTGATGGGCAAGTCGTCGCAACAAATCATTTGAAATTTGCG 1992
QY      535 TTCACTCAACAACCTTGGAGTTTTCGACGCGAGTT 570
Db      1991 AGCTTTAGAACGTCTGGAATTTACTGACCAAGAATT 1956

RESULT 3
LOCUS   AF075981
DEFINITION
AF075981 Salmonella typhimurium LT2, Lambda DASH II Salmonella
typhimurium genomic clone 390-T3, genomic survey sequence.
ACCESSION
AF075981
VERSION
AF075981.1 GI:3320851
KEYWORDS
GSS.
SOURCE
Salmonella typhimurium
ORGANISM
Salmonella typhimurium
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
REFERENCE
1 (bases 1 to 860)
Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.
Sample sequencing of a Salmonella typhimurium LT2 lambda library:
comparison to the Escherichia coli K12 genome
FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
9243757
10227170
Contact: McClelland M
Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA
Email: mclelland@lifsci.sdsu.edu
Class: shotgun.
FEATURES
Location/Qualifiers
source
1..860
/organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="LT2"
/db_xref="taxon:602"
/clone="390-T3"
/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"

ORIGIN
Query Match 3.0%; Score 72; DB 28; Length 860;
Best Local Similarity 54.7%; Pred. No. 9.9e-08;
Matches 220; Conservative 0; Mismatches 160; Indels 22; Gaps 3;

QY 52 TGAAGAAGGCTTTTACGTGGTATTCTTCTACGGTCCAGAGCTCACAGCGAGCGGC 111
Db 18 TGGCAAGCGTTGTACGTGGGATCTTAACATCTCTGCGGATCTGGCCAGACAGCTAT 77

QY 112 TGAGTTATCGCGGAGAGGGCTGCGGCTTCTGATCATCAGCCAGCTATTCCATCAT 171
Db 78 CGATATCTCGGAGATCTCGGACGCGCTTGCCTGATCATCAGCCATAATTTTCGCTTT 137

QY 172 TAATCTTGGGTGGAGAACCGGGCGATGACGGGTGAGAACTTTGTCAGTCACTGCCAA 231
Db 138 TGAGCGTTGGGTAGAGAGCGGGC-----TGCTGGGCTGTGTGAGA 179

QY 232 CAATGCTCTTGGCGTCATTGCTTCTCACTTCACATTGCGCGAGGCGCTGCTCAAGCAATA 291
Db 180 AAAAAGCGTCGTAGTATTGCCCTTCTCGCGCTGGGGGCGGCA-GCTCACTGACCGTAA 238

QY 292 TCTCGATGATTCACAGAGGTTCCGCG---GCCAGCGAGGTTAAGTCCCTCTGTGAGGG 348
Db 239 TTTGAATGGTATTCGGAAGATTCCCGCGGCGGAGCGGAGCGGTTTCTTCTTAACACAGA 298

QY 349 CATGTTGAACGTGAACAATATGATATGCTCCGCAAGCTCAATGACATCGCCCGACGACG 408
Db 299 ACAGATTACCGCGCAAACTGAAAAAGTTGCGCGGTTGGAATGACTGCTCGCGAGCG 358

```

```

QY 409 CGGCGAGTCACCTGGCGAGATGGCGCTTGCATGGGTGCTGCG 450
Db 359 GGGTCAAAATTTGATCAATGGCGCTCGCTGGATGCTGCG 400

RESULT 4
LOCUS   CNS01MQH
DEFINITION
CNS01MQH 830 bp DNA linear GSS 14-JUN-2001
Anopheles gambiae GSS T7 end of clone 22E24 of NotreDamel library
from strain PRST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION
AL151258
VERSION
AL151258.1 GI:7011737
KEYWORDS
GSS.
SOURCE
Anopheles gambiae (African malaria mosquito)
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;
Anopheles.
REFERENCE
1 (bases 1 to 830)
Roth, C.W., Brey, P.T., Ke, Z., Collins, F.H. and Weissenbach, J.
Genoscope.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
REFERENCE
2 (bases 1 to 830)
Roth, C.W., Brey, P.T., Ke, Z., Collins, F.H. and Weissenbach, J.
Direct Submission
Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
COMMENT
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
Pasteur.
FEATURES
Location/Qualifiers
source
1..830
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PEST"
/db_xref="taxon:7165"
/clone="22E24"
/clone_lib="NotreDamel"
/note="end : T7"

ORIGIN
Query Match 2.5%; Score 60.4; DB 29; Length 830;
Best Local Similarity 50.7%; Pred. No. 0.00013;
Matches 142; Conservative 1; Mismatches 137; Indels 0; Gaps 0;

QY 1058 GCGGCTTCCACCTTTGTTGTTCATGGCGTCTTTCGTCGCCATGACGGCAACATAACAGG 1117
Db 318 GCCAGTTTCAGGTTACTGCTCATCGCGGCTTTTCAGCGCGCGGAATCCGTACAGAGCAGA 377

QY 1118 TAAGCATCCACCCAGCCGATATATCGAGCAGCATCGGCGGCGATTGGACAAAGA 1177
Db 378 AACGCCACGCCGCCCGCGAGTGACCAACGCCAGCAGCAGCAGATGCGATCAGCGCGG 437

QY 1178 TCACGCCCAAGGTGCGCGCGATGACAAAGACGTCAGAAATTAACACACAGAGAAGA 1237
Db 438 CTGCCGCCAAATCCCGGACAGATCAGCAGCAATCACTTACCGCGCATAGCGGCA 497

QY 1238 ACCGCAATGAGTCTTCGCGCTTAATTCCTTGTGTTAATCACAGTACATTCGCGGTCG 1297
Db 498 ATCATCAGATGCTACTGCGCGCGATGCGCTGATTATCATCAGACGCAATTTTGGCGCGG 557

QY 1298 ATGGAAGTAAAGACTGGCCCCCAAGAGCAGCTGTA 1337
Db 558 AGGGAAGAAATAAAAAATGACCTAAGGCAAGCCCTTGAM 597

RESULT 5
LOCUS   B2554294
DEFINITION
B2554294 1034 bp DNA linear GSS 17-DEC-2002

```

AUTHORS Spencer,D.H., Raymond,C.K.Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.

TITLE Whole-Genome-Sequence variation among multiple isolates of *Pseudomonas aeruginosa* library

JOURNAL J. Bacteriol. (2002) In press

COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source	1. 897
	Location/Qualifiers
	/organism="Pseudomonas aeruginosa"
	/mol_type="genomic DNA"
	/strain="2-164"
	/db_xref="taxon:287"
	/clone="pacs2-164_2569"
	/clone_lib="pacs2-164"
	/note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

Query Match	2.3%;	Score 54.2;	DB 28;	Length 897;
Best Local Similarity	52.8%;	Pred. No. 0.0067;		
Matches	143;	Conservative 0;	Mismatches 120;	Indels 8; Gaps 1;

QY 650 TAACCCATCACATCATGTTTATGCGCCATCGGTGATCATCACTGCCACGACGAGCTTG 709

DB 290 TATCCCGGAACAGCAGTGTGGCGGCCATGCCCCAGCATCATGCGCCGACCATCAGTCCG 349

QY 710 ATCCAGCGCCACACCTTGGGCTGTGAGCAGCGGGCGTGACAAATGCTGCTGGCCGAAACCC 769

DB 350 AGCAGCGCCAGTGGCGCGCGCCAGCAGGSGGCGAGCCATGCTCGCGCGGAGGCG 409

QY 770 ACCAGCGGGAACACAGATCAGCTTTGCCGGGACCGGGCGAGCGCGGAAATCCACCGTCGG 829

DB 410 AGGCGGAAGAACACATCAGCAGCGCTTTGGCGCGCGCGAGGCGCATNGCGCCCGG 465

QY 830 GTGTGCGGTATTGCGCGCGCGCGCGCGATAAACAACGCGTCCAAATACGATTC 889

DB 466 ----CGCGNCTTGTGGCGCGCGAGCGTAAAGTAGCTACGTATCGATAGACGTGG 521

QY 890 GGGTTCAACGAGTCAAGCAGCATTTGCCATCA 920

DB 522 GGGTTAAGCCAGTGAAGCCAGTGGCCGCA 552

RESULT 7

BZ568946

LOCUS

DEFINITION

pac2-164_8165.y2 pac2-164 *Pseudomonas aeruginosa* genomic clone

pac2-164_8165, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Pseudomonas aeruginosa
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; *Pseudomonadales*;
Pseudomonadaceae; *Pseudomonas*.
1 (bases 1 to 1620)

REFERENCE

AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.

TITLE Whole-Genome-Sequence variation among multiple isolates of *Pseudomonas aeruginosa* library

JOURNAL J. Bacteriol. (2002) In press

COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954

[illegible][illegible]


```

Db      602 MATOMMMKMTKMTAAAMMMMMMMMMMMMMMMMMMMGAAAGMMMMMMMMMAAGMMGM 543
QY      1176 GATCAACGCCCAAGTGGCGGATGAACAAAGAGAGTCCAGAAATTAACACACAGAGAA 1235
Db      542 MGVAAAGMMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGM 483
QY      1236 GAACCGCAATAGTCTTTCGGCTTAAATTCCTTTTAAATCACCAGTACATCTTCGGCT 1295
Db      482 MMMAAAMMMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGMGM 423
QY      1296 CGATGACAGTAAAGACTGCCGCCCAAGAGACACCTGTAA 1337
Db      422 AAAAAAMCAAMMAGMARHMMCCMMARAAAMVAVCMGAAA 381

RESULT 12
LOCUS   CA830247
DEFINITION 1117004E09.y1 1117 - Unigene V from Maize Genome Project Zea mays
cDNA, mRNA sequence.
ACCESSION CA830247
VERSION   CA830247.1 GI:26558012
KEYWORDS EST.
SOURCE   Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
TITLE     clade; Panicoidae; Andropogoneae; Zea.
JOURNAL   1 (bases 1 to 614)
COMMENT   Walbot V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1117004 row: E column: 09.
FEATURES
source
1..614
Location/Qualifiers
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="dbEST:946110C04.y1"
/db_xref="taxon:4577"
/clone_lib="1117 - Unigene V from Maize Genome Project"
/note="This library represents the unique genes found in
the fifth round of EST sequencing at Stanford University
for the maize genome project. Sequences are present from
library 946. Contigs were assembled using ZmDBAssembler
and 2 representatives from each contig were selected for
the Unigene set. All singlets were also selected."

ORIGIN
Query Match 1.9%; Score 46; DB 14; Length 614;
Best Local Similarity 46.3%; Pred. No. 0.9;
Matches 151; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY      686 ATCAACATGCGCAGCAGCTGTGATCCAGCGCACACCTTGGGGCTGGACAGCGGGCT 745
Db      7 ATCCCAATATCCGAACGGCAATGTCTCGCTGCCCGCATCGTCTCTCGCGCTCGCC 66
QY      746 GACAATGCTGTGCGCGCAAAACCCACAGCGCGAAACAGATCAGGCTTGGCGGAACGG 805
Db      67 GCCGCGCGCGCGAGCGGTTCAGCGGCTCAGCCAGCGGCTCAGCGCGGGCGGAGACGAC 126
QY      806 CAGCGCGGCAAAATCCACCGTCCGGTGTGCGCGTATTCGCGCGCGAGCGCGCGATTAAC 865
Db      127 CTACTCCCAAGTACGGCTCCCGAAGGGGCTCATCCGAGCTCCGTCGCTCTCTACAGC 186

```

```

QY      866 ACAAACCGCTCCAAATACGCATTCGGGTTCAACACAGGTCCAGCAGATTCGATCAACATG 925
Db      187 TTGACGAGGCGCACGCGCGCTTCGAGATCCACCTCGCGGACCTGCTACGCTCCACTTC 246
QY      926 GCGTTTACCAACACCGCTGCTTATCGAGCTCAGCTCCACCCGCGACCGGTTGGGGTG 985
Db      247 GCGTCCCGACCTCGTCTACTCAGAGAGGACCATTAACCGCAAGCTCTCCAAAGGCGGCCATC 306
QY      986 TCAGTGGCCACCGCGCGAACCAGCCCAA 1011
Db      307 TGGACCTCTCCGGGCTCAGGCCAA 332

RESULT 13
LOCUS   CD881613/c
DEFINITION F1.103L23F010329 F1 Triticum aestivum cDNA clone F1103L23, mRNA
sequence.
ACCESSION CD881613
VERSION   CD881613.1 GI:32641286
KEYWORDS EST.
SOURCE   Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
TITLE     Pooidae; Triticeae; Triticum.
JOURNAL   1 (bases 1 to 720)
COMMENT   Genoplatne, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplatne
Genoplatne
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplatne' (http://www.genoplatne.com
and http://genoplatne-info.inbioigen.fr).
FEATURES
source
1..720
Location/Qualifiers
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="F1103L23"
/tissue_type="leaf one"
/clone_lib="F1"

ORIGIN
Query Match 1.9%; Score 46; DB 14; Length 720;
Best Local Similarity 48.1%; Pred. No. 0.98;
Matches 130; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY      686 ATCAACATGCGCAGCAGCTGTGATCCAGCGCACACCTTGGGGCTGGACAGCGGGCT 745
Db      520 ATGCCACAGCTCGCTCCCGAGGCTTTTCAGAGCTCCAGCTTCGGGGTGTGGCGTGGCC 461
QY      746 GACAATGCTGTGCGCGCAAAACCCACAGCGCGAAACAGATCAGGCTTGGCGGAACGG 805
Db      460 GCCACGCTCGACGCGCGCTAAACTCTCTTCGCCAGCTGGATGACGAGGGTCCCGACTCG 401
QY      806 CAGCGCGCAAAATCCACCGTCCGGTGTGCGCGTATTCGCGCGCGAGCGCGCGATTAAC 865
Db      400 CCGCGCGCGCGAGGAGCAGGATGACATTTGCGCGGAGAGCGCGCTCTCTCGAGCGCC 341
QY      866 ACAAACCGCTCCAAATACGCATTCGGGTTCAACACAGGTCCAGCAGATTCGATCAACATG 925
Db      340 TCATGGCGGCTCTCGATGGCGAGCGGCGAGACGGCGGCTTGGCGAGTCTGATCTCTTG 281
QY      926 GCGTTTACCAACCGCGCTCTTATCGAGC 955
Db      280 GGCTTTGAGCGCGCACCGACTTCTCTCCAGC 251

```


Job time : 6529.5 secs